A HCl technique for improving requirements elicitation

Silvia T. Acuña a, John W. Castro a,*, Natalia Juristo b

a Departamento de Ingeniería Informática, Universidad Autónoma de Madrid, Calle Francisco Tomás y Valiente 11, 28049 Madrid, Spain
b Facultad de Informática, Universidad Politécnica de Madrid, Campus de Montegancedo s/n, 28660 Boadilla del Monte, Madrid, Spain

1. Introduction

HCl as “a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use...” [1]. This involves studying and getting acquainted with people as members of groups or organizations, the conditions under which subjects are likely to want to use their device, as well as the characteristics involved in this interaction. HCI processes and techniques assure that the output software product conforms to minimum usability standards.

Pure HCl techniques usually have two shortcomings that restrain their use within SE processes [2–4], although there are sound proposals for integrating usability engineering proposals with SE processes [5,6]. On the one hand, they do not have a well-defined, complete and detailed procedure to guide the software engineer through the proper application of the technique. On the other hand, they do not prescribe products output by applying the techniques.

The Personas technique [7] gathers, analyses and synthesizes information related to the users that are to interact with the software system. It helps to focus software analysis and design on end user features and goals. However, it shares the above shortcomings of HCl techniques: it has no detailed definition of activities and products. These problems make the introduction of Personas into the SE requirements stage overly complex and unclear for developers.

According to HCI, there should be an understanding not only of users’ needs and goals but also of their characteristics and capabilities in order to design and implement a usable system [8]. The understanding of the people that interact with the system should constitute the groundwork for software development. The SE requirements activity could be improved by incorporating Personas techniques to understand the user. The goal of our research is to modify Personas to readily build the technique into the requirements stage of regular SE developments.
the impact of building Personas into the requirements stage. In Section 6 we discuss an application example analysing the feasibility of our proposal. Finally, Section 7 addresses the conclusions.

2. Related work

Personas is attributed to Cooper [9], who later updated the method in [2,7]. On this groundwork, several versions [10–13] have been proposed for successfully generating personas, the main final product output by the technique. However, some descriptions on how to apply the Personas technique are fuzzy and unordered, and results are unrelated [2,7,14,18]. To assure that users’ viewpoints are the focus during software design, the Personas technique creates fictitious users called personas rather than taking into account real users. Personas characterize the target users on whom development efforts should focus. The main benefit of Personas is to determine whom the development should target.

Personas perform a survey of users that is used to pair personas’ characteristics and goals. Goodwin [10] suggests that personas should be primarily based on qualitative data that are gathered from interviews and observations. Quantitative data can be used, time and budget permitting, to validate the results of the qualitative research. Cooper and Reimann [2] share Goodwin’s view and further detail the social research methods that can be used to focus more on user goals than on their tasks and take into account the customer domain.

The data gathered from observations and interviews are mapped to behavioural variables. Behavioural variables are types of behaviour that range across a spectrum [7,19]. Cooper et al. [7] suggests identifying behavioural variables of the type: activities, attitudes, aptitudes, motivations and abilities. Aquino and Leite [20] and Pruitt and Grudin [14] state that the mapping aims to show where each interview subject is positioned with respect to the selected behavioural variables.

A set of subjects who cluster around 6–12 different variables will likely represent a significant behaviour pattern that will form the basis of a persona [7]. A persona is built by adding details to the behavioural pattern. For a pattern to be valid, there must be a logical or causal relation between the grouped behaviours [10] rather than a spurious or spontaneous correlation. For example, there is a clear logical relation if the data show that people that regularly buy CDs also like to download MP3 files; but there is unlikely to be any clear logical relation if the data show that people that regularly buy CDs also like to collect stamps.

Once the personas have been created, they need to be documented and communicated to the development team. Pruitt and Grudin [14] identify the communication of personas as a key success factor. In a failed application of Personas, reported by Blomquist and Arvola [15], communication breakdown was identified as the key reason for failure. To prevent this, Cooper and Reimann [2] mention two basic deliverables for each created persona: a list of characteristics and a third-person narrative about the persona. This list of characteristics is not just a specification of brief points describing behavioural characteristics but it also needs to include observed possible behaviours. However, an overly fabricated life story, too many quirks, etc., are a distraction and make the personas less credible [21]. Factual data are the only sound basis for the development and business decisions to be taken by the development team [7,8].

Cooper and Reimann stress the importance of the persona having a name and a photograph for realism. The narrative should be one or two pages long and not cover all the observed details, as the team members will ideally have participated in the research phase, and team non-members do not need to know the details of the research [2]. The best narrative is one that briefly introduces the persona in terms of his or her job or lifestyle and concisely outlines a day in his or her life, including concerns and interests that have direct implications for the product [8]. When the personas have been documented and materials are complete, a meeting should be arranged with the development team to introduce and present the personas [14].

We have analysed the major versions of the Personas technique referenced in the literature [7,8,14,18]. We have focused on how the authors describe the procedure to be enacted to apply the technique and the resulting documents. Generally, Cooper’s Personas technique [7] explains the technique procedure, that is, what is does, although some steps are not completely defined. As regards the formalization of the products resulting from applying the technique, Cooper only explicitly mentions two products: list of characteristics and relevant goals and a narrative.

On the other hand, Pruitt and Adlin’s version of the Personas technique [8] gives a full description of what Personas are, together with instructions on how to apply the technique. This version of the technique was originally inspired by Cooper and Reimann [2]. Unlike Cooper’s technique, Pruitt and Adlin’s version sets out a procedure that should be enacted as a five-phase life cycle (familly planning, conception and gestation, birth and maturation, life-time achievement and retirement), each of which explicitly describes what to do and defines how to do it. Pruitt and Adlin establishes a set of deliverables for each personas life cycle phase, which Cooper does not do [7]. Pruitt and Adlin introduce examples for half of these documents, although they only give a formalized structure for the Personas Foundation Document.

Pruitt and Grudin’s version of the Personas technique [14] is based on Cooper’s Personas technique [7]. Pruitt and Grudin actively used and refined the technique for about 3 years. Pruitt and Grudin’s adaptation of the Personas technique gives a general description of the procedure to be enacted to apply the technique which is not explicitly divided into steps like Cooper’s technique [7]. Pruitt and Grudin use a series of panels to explain this procedure. The panels narrate what to do to apply the technique, Pruitt and Grudin explicitly mention the following documents: narrative, foundation, posters/flyers and a feature-persona weighted priority matrix. Like Cooper, Pruitt and Grudin also define the Foundation and Narrative Documents, although of these documents are not formalized (they do not explicitly specify the document sections).

Aoyama [18] proposes the Persona-Scenario-Goal (PSG) methodology as a user-centred goal-oriented requirements engineering methodology. This methodology helps with requirements elicitation in domains where applications have many users (for example, Web applications). Rather than developing a version of Personas, Aoyama uses the technique as part of his PSG methodology. Aoyama gives a very general and shallow description of the procedure to be enacted to apply the technique and no more than a template for describing personas with only a general outline of its contents. This template is equivalent to the Personas Foundation Document proposed by Cooper et al. [7] and Pruitt and Adlin [8], except that it contains much less information and is really a synopsis or shell of the Foundation Document.

3. The Personas technique: strengths and weaknesses

Personas is a powerful and multi-purpose analysis technique that can help software engineers to identify the functionalities of the software system. Both the personas’ goals and tasks are essential for this purpose [7]. Developers that design for particular personas, even if they are fictional as in the case of personas:

- Design for the end user, ensuring that they do not project their own goals, motivations, skills and mental models on the software system development.
Table 1
Limitations of the Personas technique.

<table>
<thead>
<tr>
<th>Activities according to [7]</th>
<th>Limitations with respect to</th>
<th>How to perform the activity</th>
<th>Product to be output by the activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify Behavioural Variables</td>
<td>Assumed to have already been completed and that the data gathered have been roughly organized</td>
<td>No specification of how to record the hypotheses in a proper document</td>
<td>No mention of the content and structure of the product output by this activity</td>
</tr>
<tr>
<td>Map Interview Subjects to Behavioural Variables</td>
<td>No indication of how to output the ranges of the key behavioural variables for mapping</td>
<td>No formal specification of the content and structure of the product output by this activity</td>
<td>No formal specification of the content of the product and no specification of the structure at all</td>
</tr>
<tr>
<td>Identify Significant Behavioural Patterns Synthesize Characteristics and Relevant Goals</td>
<td>No explicit instructions on how to output these patterns</td>
<td>No explicit specification of the procedure to be enacted to carry out this activity</td>
<td>No mention of the content and structure of the product output by this activity</td>
</tr>
<tr>
<td>Check for Completeness and Redundancy</td>
<td>No specification of how to synthesize characteristics</td>
<td>No mention of the content and structure of the product output by this activity</td>
<td>No mention of the content and structure of the product output by this activity</td>
</tr>
<tr>
<td>Expand the Description of Attributes and Behaviours</td>
<td>No explicit specification of the procedure to be enacted to carry out this activity</td>
<td>No mention of the content and structure of the product output by this activity</td>
<td>No mention of the content and structure of the product output by this activity</td>
</tr>
<tr>
<td>Designate Persona Types</td>
<td>No specification of any product output by this step</td>
<td>No mention of the content and structure of the product output by this activity</td>
<td>No mention of the content and structure of the product output by this activity</td>
</tr>
</tbody>
</table>

Table 2
Set of activities proposed for Personas (new activities are italicized).

<table>
<thead>
<tr>
<th>Activities</th>
<th>References discussing similar activities</th>
<th>Objectives</th>
<th>Products</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTIVITY 1: STATE HYPOTHESES</td>
<td>Activity 1.1: Identify Possible Personas</td>
<td>State preliminary hypotheses about the possible personas to be created</td>
<td>List of hypotheses for Personas</td>
</tr>
<tr>
<td></td>
<td>Activity 1.2: Hold Ethnographic Interviews</td>
<td>Based on these hypotheses, investigate possible system users to find out their motivations and behaviours, gathering behavioural data</td>
<td>Transcribed interviews</td>
</tr>
<tr>
<td>ACTIVITY 2: IDENTIFY BEHAVIOURAL VARIABLES</td>
<td>Activity 2.1: Synthesize Interview Responses</td>
<td>Synthesize the responses to all the interviews</td>
<td>List of behavioural variables</td>
</tr>
<tr>
<td></td>
<td>Activity 2.2: List Behavioural Variables</td>
<td>List all behavioural variables. Check identified hypotheses for validity.</td>
<td>Interview synthesis</td>
</tr>
<tr>
<td>ACTIVITY 3: MAP INTERVIEW SUBJECTS TO BEHAVIOURAL VARIABLES</td>
<td>Activity 3.1. Identify the Ranges of Behavioural Variable Values</td>
<td>For each behavioural variable identify its range of possible values</td>
<td>Ranges of Behavioural Variables</td>
</tr>
<tr>
<td></td>
<td>Activity 3.2: Map Interview Subjects to Behavioural Variables</td>
<td>Represent exactly how the multiple subjects are grouped with respect to each of the significant behavioural variables</td>
<td>Mapping of Interview Subjects</td>
</tr>
<tr>
<td>ACTIVITY 4: IDENTIFY SIGNIFICANT BEHAVIOUR PATTERNS</td>
<td>Activity 8.1: Select Represent alive Personas to Elicit Requirements</td>
<td>Prioritize the created personas to determine which should be the primary design objective: that is, find just one primary persona whose needs and objectives can be completely and positively satisfied by a single interface</td>
<td>Persona Type Association</td>
</tr>
<tr>
<td></td>
<td>Activity 8.2: Enrich the System with Secondary Personal</td>
<td>Determine what secondary persona needs are likely to enrich the system</td>
<td>(Software Requirements Specification is enriched)</td>
</tr>
<tr>
<td>ACTIVITY 5: SYNTHESIZE CHARACTERISTICS AND RELEVANT GOALS</td>
<td>[7,8,14]</td>
<td>Synthesize characteristics and relevant goals. Describe the personas' personalities</td>
<td>Personas Foundation Document</td>
</tr>
<tr>
<td>ACTIVITY 6: CHECK FOR REDUNDANCY AND COMPLETENESS</td>
<td>[7]</td>
<td>Check persona mappings, characteristics and goals</td>
<td>Validation Document</td>
</tr>
<tr>
<td>ACTIVITY 7: EXPAND THE DESCRIPTION OF ATTRIBUTES AND BEHAVIOURS</td>
<td>[7,8]</td>
<td>Convey the attitudes, personality, needs and problems of the personas to other team members</td>
<td>Narrative</td>
</tr>
<tr>
<td>ACTIVITY 8: DESIGNATE PERSONA TYPES</td>
<td>Activity 10.1: Implement Mock-Ups</td>
<td>Build annotated use case diagram and draft use case specification</td>
<td>Annotated Use Case Diagram</td>
</tr>
<tr>
<td></td>
<td>Activity 10.2: Evaluate Mock-Ups</td>
<td>Validate mock-ups</td>
<td>Use Case Specification</td>
</tr>
<tr>
<td>ACTIVITY 9: BUILD USE CASES</td>
<td></td>
<td></td>
<td>Mock-ups</td>
</tr>
<tr>
<td>ACTIVITY 10: IMPLEMENT AND EVALUATE PROTOTYPES</td>
<td></td>
<td></td>
<td>Evaluation Document</td>
</tr>
</tbody>
</table>
• Prioritize functionalities, ensuring that they do not build functionalities that future users might never use. For example, the developer might wonder, “Does Sandra Wolf need to do this operation very often?” The response to questions like this helps to prioritize the functions that the software system should have according to its users.

We tried to apply Cooper’s version of the Personas technique [7] and we came up against the following obstacles.

The first step of the technique recommended by Cooper is Identify Behavioural Variables, that is, he assumes that user research has already been completed and the gathered data have been roughly organized. But the technique does not specifically mention this task. The user survey necessary for eliciting behavioural variables is an implicit activity rather than being stated as the first step of the technique.

Applying Cooper’s Step 2 (Map Interview Subjects to Behavioural Variables), we found that he does not indicate how to output the behavioural variable ranges necessary for mapping subjects. Additionally, there is no specification in Step 2 of the product to be generated as the output of the step or how the information should be structured.

In Step 3 (Identify Significant Behaviour Patterns) there is no explicit specification of how these patterns are output; Cooper states that subjects should be grouped in the previously output mapping. As in the earlier steps, there is no specification of the resulting product or of what its content or structure should be like.

We came across other similar limitations (up to a total of 11) along the lines of the above three examples. Table 1 shows the limitations that a software engineer encounters when trying to apply Personas. In each of the steps, we found shortcomings in both the definition of the procedure to be enacted and the formalization of the product resulting from the execution of each step.

4. Aligning Personas with SE

For each of the limitations found in Table 1, we devised an improvement to be built into Personas. We have incorporated these improvements into a SE version of Personas. The improved Personas avoid the weaknesses encountered by an average software developer unfamiliar with HCI techniques applying the original Personas. This adaptation of the Personas technique is useful for organizations that have a defined and formalized process using non-agile practices, as Personas involves heavy weight practices like interviewing users, synthesizing interviews and mapping interview subjects.

To be able to build Personas into routine SE developments, it is necessary to define activities and products associated with each activity. We opted to incorporate these improvements into the latest version of the Personas technique published by Cooper et al. [7]. The grounds for this choice were: (i) Cooper made the original proposal; (ii) this proposal was the groundwork for research by other authors; and (iii) this proposal has been successfully used in a number of real projects [11,13,16,17].

Our proposal is composed of a group of activities that, together, lead to the creation of Personas. Table 2 sets out these activities with their objectives, techniques and associated products. The new activities or products (not included in any other earlier version of Personas) are italicized in Table 2. Note that, even for existing activities, it has been necessary to identify and define both the actual activities and their products using all the existing versions of Personas as sources.

In the new first activity, State Hypotheses for Personas, we propose to generate a List of Hypotheses for the Personas 1 that are to be created, as well as developing and holding the interviews with potential users, taking the responses from the Transcribed Interviews to then gather the information required to carry out other activities. The aim behind the hypotheses for Personas is to identify variables that can make a difference between users based on their needs and behaviours. Table 3 shows an example of a List of Hypotheses for Personas. The identification of hypotheses for Personas is the first activity to be performed to identify and synthesize personas.

As part of the activity Identify Behavioural Variables, we propose a new activity for synthesizing each response to the interviews held in the earlier activity as behavioural variables. Cooper et al. [7] recommend that a list of behavioural variables should be generated from the different aspects of the observed behaviours (Activity 2.2) but this is hard to do from a disaggregated set of information like a set of interviews. For this reason, we add an intermediate step that synthesizes the different interviews into common behaviours.

Developers need to elicit the key behavioural variables for every software system as they are domain and task dependent. The application domain is one of the aspects considered in interview design. Often the behavioural variables are implicit in each of the hypotheses for Personas created in the initial activity where the research was conducted and information gathered. For example, if one of the hypotheses is for a university student, it could be assumed that he will obviously be aged between 18 and 25 years and have some knowledge of computers, although this needs to be double-checked. The implicit variables in this case then are age and computer literacy. Some examples of the List of Behavioural Variables for an e-commerce application are frequency of shopping, desire to shop, motivation to shop and age.

As part of the activity Map Interview Subjects to Behavioural Variables, we propose a new activity for identifying the possible value ranges of each behavioural variable. For example, the possible values for the behavioural variable Frequency of Shopping range from frequent to infrequent. We propose that these variable value ranges should be identified by synthesizing the interview responses. Table 4 illustrates an example of behaviour variable range values.

Interview subjects can be more readily mapped if the value ranges have been defined before the mapping. Cooper et al. [7] describes mapping as positioning each interview subject with respect to one of the behavioural values. The accuracy of this mapping is not critical, but it is important to identify where each interview subject is in relation to the others. There is often no way of measuring the accuracy of this mapping: we have to rely on the perceptiveness of the person doing the mapping. Mapping is done on the basis of the observations of the subjects. Fig. 1 shows a fragment of a mapping of interview subjects. Each subject is represented by a smiley. Each smiley has been assigned a label (a number from 1 to 5), as well as a different shade of grey.

The aim of the activity Identify Significant Behaviour Patterns [7] is to detect groupings of particular subjects around multiple ranges or values. A group of subjects clustered around six to twelve different variables will represent a significant behaviour pattern that will form the basis of a persona. We propose that this activity output two products: Percentage Grouping Table and Significant Behaviour Patterns. Table 5 illustrates an example of Percentage Grouping Table.

The percentage grouping table contains the percentage of interview subjects that have been mapped to each behavioural variable value. This percentage is calculated over the total number of interview subjects. Fig. 2 illustrates an example of Significant Behaviour Patterns.

During the activity of Synthesize Characteristics and Relevant Goals for each identified significant behaviour pattern, we propose that the details of the data gathered from interviews (State Hypotheses Activity) and from the synthesis of the interviews

---

1 Products and activities are capitalized.
(Identify Behavioural Variables Activity) be synthesized; that is, the environment, a typical work day (or other important period of time), dissatisfactions, and key relationships of potential users with others need to be described. Note that we are creating a design tool not an outline of a character for a novel. As a result of this activity, we propose to create, for each identified personas, the Personas Foundation Document. When deciding which characteristics to include in the foundation document, we should think about the type of information that will be useful to the development team. Fig. 3 shows the structure and information that we propose for the Personas Foundation Document.

The activity Check for Redundancy and Completeness is performed [7] to search for missing information and knowledge. Additional research may be required to discover behaviours that are not to be found in the key behaviour patterns. This would have an impact on all the other activities. We may also have to check the notes to see if there is any persona that needs to be added to satisfy stakeholder assumptions or requests [7]. To do this validation, a number of questions have to be answered, such as are all the created personas significantly different? Or, do all the created personas represent the diversity of real-world behaviours and needs that the system will address? If two personas appear to differ only as to sociodemographic data, either one of the redundant personas can be eliminated or the characteristics of personas should be further specified to show up the distinctions. At least one significant behaviour should distinguish one persona from any other. We have to assure that, as a whole, the personas are different and complete, and the set of personas is manageable. The proposed output of this activity is the Validation Document. Fig. 4 shows a possible structure for this document. The Validation Document includes a conclusions section specifying whether or not it is necessary to formulate new personas hypotheses and run through the technique activities again. It also specifies which documents, if any need to be modified such as, for example, interview subject mapping.

The Personas Foundation Document that we propose states the fundamental nature of complex behaviours but leaves many finer points unmentioned. For this reason, there is an Expand the Description of the Attributes and Behaviours Activity where a narrative is created for each persona [7]. The narrative is a one-page document describing the persona and a typical day in his or her life. The narrative should have a conclusion expressing what the persona is looking for in the product. Fig. 5 shows an example of a narrative.

Knowledge of the audience that is to use the software helps developers to fix a definite and explicit development objective. The created personas represent the potential audience. The main interface of the software system should be designed for system use by a persona. The created personas have to be prioritized to determine which should be the primary development target. We have to find just one of all the personas whose needs and goals can be completely and opportunely satisfied by a single interface without disappointing any of the other personas. To do this, a personas type designation process is enacted [7] taking into account all the information gathered in the earlier activities. This process is carried out in the Designate Persona Types Activity. Cooper et al. [7] proposes two main types of personas: primary and secondary. Primary personas are those that represent the primary

![Fig. 1. Fragment of the mapping of interview subjects for an e-commerce application.](image-url)
software development goal, and secondary personas are those that have some additional specific needs that the primary persona’s interface is unable to cover, although they do not interfere with the functionalities offered by the interface for the primary persona. Alternative interfaces, for example, use profiles for novice or expert users, are generated for these personas.

We have defined a new activity that links the research done using Personas with the remainder of the development process: Build Use Cases, where use cases are put together based on the Personas Foundation Document and Narrative and the user knowledge acquired throughout all the previous activities. As a result of this activity, we propose that an Annotated Use Case Diagram be built. This diagram is based on the traditional use case diagram, to which we add a brief description of each of the personas involved in the use case. We suggest that the description of the persona contain a brief outline of aspects like persona name and type, as well as an informative note on the use case.

Finally, we propose a new activity, Implement and Evaluate Prototypes, where mock-ups are built and evaluated. The use cases detailed in the last activity and the knowledge of the user acquired through the other Personas activities need to be taken into account to build the mock-ups. The evaluation that we propose is to be conducted in the actual environment where potential users routinely carry out their tasks or in convenient meeting rooms where the evaluator can meet with potential users.

5. Enriching requirements phases with Personas activities

As Personas synthesizes data about the users to help identify what the product should do to satisfy users’ needs and profile, the software requirements stage should be the best development phase for incorporating this technique. As the Personas technique aims to get acquainted with users from the viewpoint of the software development process, the technique has to be applied first. The requirements process will not start until the technique has been applied and we know who the primary user of the software system will be.

To identify where the Personas activities will have an impact, we consider SWEBOK RE activities [22]: requirements elicitation, requirements analysis, requirements specification and requirements validation. We have used SWEBOK RE as the baseline requirements process activities for the purpose of making it clear for all developers which requirements activities should include the proposed Personas issues. The right-hand side of Fig. 6 shows these four activities. Each SE activity is linked to one or more Personas technique activities (left-hand side of Fig. 6). Note that Personas activities are not assigned to the right-hand side of Fig. 6 in any particular order but according to the type of SWEBOK activity to which the Personas activity maps. For example, “Identify Possible Personas” and “Hold Ethnographic Interviews” are “Requirements Elicitation” activities. The aim is to identify the knowledge generated by the Personas technique and how this knowledge enriches the requirements process. Elicitation and requirements analysis are the activities that benefit most from this enrichment.

In the following we explain why the Personas technique activities were mapped to each of the requirements phase activities.

Regarding the Requirements Elicitation activity, Personas offers additional knowledge sources for eliciting information. The Personas activities that belong to the requirements elicitation activity are:

- Identify Possible Personas: Here hypotheses are formulated for possible personas. These hypotheses help to determine who the potential interview subjects will be. This is the first stage in getting acquainted with the user.
- **Hold Ethnographic Interviews:** Here ethnographic interviews are designed and held taking into account the formulated hypotheses for personas.

- **Synthesize Interview Responses:** Interview synthesis is based on an analysis and synthesis of the interviews for identifying behavioural variables.

- **List Behavioural Variables:** A list of behavioural variables that characterize possible users is taken from the synthesized interviews.

- **Identify the Ranges of Behavioural Variable Values:** The behavioural variable values are obtained by grouping subjects around the behavioural variables. These clusters characterize and provide better knowledge about potential system users.

- **Select Representative Personas to Elicit Requirements:** In this activity, the potential users for routine requirements elicitation are selected.

- **Implement Mock-Ups:** The goal of building mock-ups is to output discussion-based information. Mock-ups explicitly show what the user needs the system to do. The discussion of this mock-up with possible users will end up eliciting additional information.

With respect to **Requirements Analysis**, Personas offers useful complementary conceptual tools and/or extends those usually used in this activity. Personas activities are able to analyse the information about and knowledge of the users. Personas activities that enrich the requirements analysis activities are:

- **Identify the Ranges of Behavioural Variable Values:** The identification of the ranges of behavioural variable values is the basis for mapping interview subjects. This groundwork can be considered as a high-level model of the user. This Personas activity has an impact on both SE elicitation and requirements analysis.

---

### Personas Foundation Document

**1. PERSONA IDENTIFICATION**
- Name, title or short description
- Age, gender
- Quotation (highlighting a key, preferentially product-related aspect of the persona)
- Photograph or brief physical description

**2. ROLES & TASKS**
- Employer name or industry
- Job or role title
- Routine activities
- Key non-routine activities
- Challenging areas or failures
- Responsibilities
- Interactions with other people, systems, products

**3. OBJECTIVES**
- Short- and long-term goals
- Motivations
- Job-related objectives
- Product-related goals
- General (life) goals, aspirations
- Stated and non-stated product expectations

**4. SEGMENT**
- Market size and influence
- International considerations
- Accessibility considerations
- General and domain demographics
  - Income and purchasing power
  - Region or city, state, country
  - Educational attainment
  - Marital status
  - Cultural information

**5. SKILLS & KNOWLEDGE**
- General computer and/or Internet use
- Frequent use of products, product knowledge
- Years of experience
- Domain knowledge
- Training
- Special skills
- Competitor knowledge

**6. CONTEXT ENVIRONMENT**
- Equipment (network connection, browser and version, operating system)
- "Day in the life of" description
  - Work styles
  - Timeline for typical day
- Specific use location
- General work, home activities, leisure activities
- Relations with other people

**7. PERSONAL & PSYCHOLOGICAL DETAILS**
- Personality traits
- Values and attitudes (political opinions, religion)
- Fears and obstacles, pet hates
- Personal belongings (car, appliances)

---

*Fig. 3.* Proposed structure for the Personas Foundation Document.
Map Interview Subjects to Behavioural Variables: The representation of how multiple subjects are grouped with respect to significant behavioural variables is a preliminary model of the system user.

Identify Significant Behaviour Patterns: Identifying clusters of particular subjects that appear in multiple ranges leads to the creation of personas (archetypal users) and ultimately models the user.

Expand the Description of Attributes and Behaviours: The development of narratives rapidly introduces the persona in terms of their job or lifestyle and conveys the persona’s attitudes, needs and problems to development team members. From a SE perspective, this is a user model in the form of a narrative.

Enrich the System with Secondary Personas: Identifying what functionalities the secondary personas would need the system to have extends the software requirements.

The Personas activities related to Requirements Specification are:

- Synthesize Characteristics and Relevant Goals: This activity generates the Personas Foundation Document. This is a brief description of characteristics and relevant goals that reflects the personality of the created personas. This document is an end user model.

Fig. 4. Proposed structure for the Validation Document.

Fig. 5. Example of a Narrative Document.
Enrich the System with Secondary Personas: This activity determines which functionalities the secondary persona expects of the system. These are functionalities that were not stated by the primary persona. Note that the requirements for the primary persona should already be included in the requirements specification as a user embodying that persona should have been chosen as the source for the elicitation. The information elicited from that user (that is, from that primary persona) will have been included in the software model and in the specification according to traditional SE RE activities.

Build Use Cases: The annotated use case diagrams and the use case specification provide project stakeholders with a model providing a common understanding for deciding what the software system should do according to each persona.

The personas activities related to Requirements Validation are:

Check for Redundancy and Completeness: The goal of this activity is to check the mappings, characteristics and goals of the personas with the aim of finding out if there are important gaps that
need to be filled. This way, we are validating the activities performed and models built previously and modifying any products output in either textual or graphical format.

- **Evaluate Mock-Ups**: A document recording the results of the user evaluation of the mock-up built is prepared.

Summarizing, the use of the Personas technique in a typical SE RE stage will have the following benefits:

- Support a proper selection of the user or users that will be the main source of information in the elicitation activity.
- Model users through the textual representation of their attitudes, needs, jobs, lifestyles, problems, etc.

These two benefits should provide the following improvements in a regular SE RE stage:

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Personas</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>H0</td>
<td>Elderly, old-age pensioners</td>
<td>Elderly and old-age pensioners have plenty of time to travel all year round.</td>
</tr>
<tr>
<td>H1</td>
<td>Professionals, employees</td>
<td>Professionals and employees generally have small- or medium-sized income on which they can travel.</td>
</tr>
<tr>
<td>H2</td>
<td>Travel agents, businessmen</td>
<td>These people have to travel regularly on job-related grounds.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A. Identification of the Interview Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What is your full name? Elizabeth Brown</td>
</tr>
<tr>
<td>2. How old are you?</td>
</tr>
<tr>
<td>From 20 to 30</td>
</tr>
<tr>
<td>___ From 51 to 60</td>
</tr>
<tr>
<td>3. Do you use glasses or contact lenses? Yes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. Roles &amp; Tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. Name of employer and position: I used to work as a Secretary at the Complutense University of Madrid.</td>
</tr>
<tr>
<td>5. What part of the company do you work for? N/A</td>
</tr>
<tr>
<td>6. What type of work do you do? N/A</td>
</tr>
<tr>
<td>7. What is your occupational status?</td>
</tr>
<tr>
<td>___ Employer</td>
</tr>
<tr>
<td>X Pensioner</td>
</tr>
<tr>
<td>8. Type of business organization:</td>
</tr>
<tr>
<td>___ Public</td>
</tr>
<tr>
<td>9. Corporation type:</td>
</tr>
<tr>
<td>___ Business</td>
</tr>
<tr>
<td>___ Other: N/A</td>
</tr>
<tr>
<td>10. Describe your routine activities on a work day: I get up make breakfast for my husband, do household chores and wait for my husband to get home.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C. Objectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>11. What are your professional motivations? N/A</td>
</tr>
<tr>
<td>12. What are your professional goals? N/A</td>
</tr>
<tr>
<td>13. What are your general life goals and aspirations? The first goal I achieved was to buy a flat, and my biggest aspiration is to save up enough money to travel abroad and discover new cultures.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D. Application domain (business niche)</th>
</tr>
</thead>
<tbody>
<tr>
<td>14. What type of documents or other information would you like the flight search engine to deliver? I would like the search engine to display possible flights, rates and times for travelling to the specified destination on one page, as I find it very time consuming to have to visit several pages to find out all the details that I need. This would simplify my search.</td>
</tr>
<tr>
<td>15. What information would you need the flight search engine to display? It would be useful if it displayed the flight details, rates, tourist cities, available hotels and any instructions on cancellations or how to extend my length of stay all on one page.</td>
</tr>
</tbody>
</table>

---

**Table 6**

List of hypotheses for Personas for the Web Flight Booking System.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Personas</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>H0</td>
<td>Elderly, old-age pensioners</td>
<td>Elderly and old-age pensioners have plenty of time to travel all year round.</td>
</tr>
<tr>
<td>H1</td>
<td>Professionals, employees</td>
<td>Professionals and employees generally have small- or medium-sized income on which they can travel.</td>
</tr>
<tr>
<td>H2</td>
<td>Travel agents, businessmen</td>
<td>These people have to travel regularly on job-related grounds.</td>
</tr>
</tbody>
</table>

**Fig. 7.** Fragment of an interview for the “Old-Age Pensioner” Persona of the Web Flight Booking System.
Elicited information will match user needs thanks to a proper selection of the elicitation source.

Design decisions will satisfy user needs thanks to proper user modelling.

6. Application example

We have applied the proposed improved version of the Personas technique to design a Web Flight Booking System that searches flights based on the user’s choice of dates, the places from and to which the passengers intend to travel, and the number of passengers.

Based on the information gained from the stakeholders, domain experts and related literature we identified the hypotheses for the personas. We recommend brainstorming, followed by a voting round at the end of the session to determine the most creative and feasible hypotheses, as the best technique for classifying and selecting hypotheses. Table 6 shows the List of Selected Hypotheses.

After generating the hypotheses we designed, held and transcribed interviews. The interviews for each hypothesis are conducted based on business domain knowledge and through the proposed ethnographic interviews template. Our proposal of
The ethnographic interview is composed of questions grouped into the following sections: Identification of the Interview Subject, Roles & Tasks, Objectives, Application Domain (business niche), Socio-demographic Questions, Skills and Knowledge, Context/Environment and Personality Factors. This division was chosen to improve later analysis and synthesis of the gathered information. For example, *How old are you?* is part of the Identification of the Interview Subject as shown in Fig. 7. Having an ethnographic interview template divided into different sections helps system developers to prepare the Personas Foundation Document, as each interview section generally maps to a section of the Foundation Document. According to the hypotheses, a total of six interviews were held, two for each of the hypotheses. Fig. 7 illustrates a fragment of one of these interviews for the hypothetical persona of “Old-Age Pensioner”.

After completing and transcribing the interviews of the subjects identified in the personas hypotheses, we synthesize each response taking into account all interview subjects responses in Activity 2 – Identify Behavioural Variables. This Synthesis will be useful for later

Fig. 8. Mapping of interview subjects for the Web Flight Booking System.
identifying the Behavioural variable ranges. We recommend the use of Atlas.ti (http://www.atlasti.com/) to complete the synthesis process. Table 7 shows a fragment of a synthesis of the responses to the interviews. For example, the results of synthesizing responses to How old are you? is that half of interview subjects are aged from 31 to 40 years, whereas the other half are aged over 60. As this question conditions the behaviour of interview subjects, it is a behavioural variable as a person aged 35 years does not interact with a software system in the same manner as person aged over 60.

After synthesizing the responses to the interview questions, we gather the behavioural variables. Behavioural variables are selected at participatory meetings. These variables are then compared with the personas hypotheses to validate these hypotheses. The behavioural variables identified in this activity are dependent on the domain of the product to be developed. For example, the “Enjoyment of Travel” behavioural variable may be a major consideration in the development of a flight booking system, but absolutely irrelevant for an online bookstore software system. Table 8 illustrates the left column of the List of Behavioural Variables elicited for this application example. Continuing with our age-related example the possible values for this variable are from 20 to 30, from 31 to 40, from 41 to 50, from 51 to 60 and over 61. These values are taken from the interview held.

After identifying all the behavioural variables revealed by the subject interviews, we go onto identify the value ranges of the behavioural variables and then map interview subjects to those variables in Activity 3 – Map Interview Subjects to Behavioural Variables. At a participatory meeting, we analyse the interview synthesis to identify the ranges of each behavioural variable. Some of these variables represent a continuous range of behaviours (for example, from novice to expert computer user), and a few will represent multiple discrete options (for example, use desktop computer vs. mobile phone). As a result of identifying the ranges for each behavioural variable, we output the Behavioural Variables Value Ranges. Table 8 shows the ranges for the behavioural variables identified in Activity 2 in the right column.

We then mapped each interview subject to each of the behavioural variables. This mapping aims to show how each interview subject is positioned with regard to the selected behavioural variables. Interview subjects are mapped on each behavioural scale. There is a behavioural scale for each defined behavioural variable. The exact position of an individual subject on a scale is less important than its relative position to other subjects. Fig. 8 illustrates the mapping of the six interview subjects. Each of the six interview subjects is represented by a smiley. Each smiley has been assigned a label (a number from 1 to 6), as well a different shade of grey. This has enabled the visual identification of significant behaviour patterns. For example, we find in Fig. 8 that six interview subjects have been mapped according the age variable into two groups: from 31 to 40 years (subjects 2, 3 and 6) and over 61 years (subjects 1, 4 and 5).

After mapping the interview subjects to the behavioural variables, we then observe the groupings of individual subjects around multiple ranges or variables in Activity 4 – Identify Significant Behavioural Patterns. We have examined the mappings of interview subjects from Activity 3 and built a table showing the percentage of interview subjects. This table contains the percentage of interview subjects that have been mapped to each of the behavioural variable values over the total number of interview subjects. Table 9 shows the table generated in this activity. Continuing with our age-related behavioural variable, Table 9 shows that 50% of interview subjects are aged between 31 and 41 years and the other 50% are aged over 60 years. This information is gathered from the previous mapping and also observed early on during the interview synthesis.

Based on the above mapping (Fig. 8), we identify significant behaviour patterns with the help of the grouping table (Table 9). Having identified these patterns, we cluster the individual subjects around multiple ranges of variable values. Fig. 9 shows the result of clustering, which led to the creation of two personas, Sandra Wolf and John Smith. These clusters are identified with the help of the percentage grouping table. The percentage grouping table is useful for identifying which two values associated with each persona type (primary and secondary) have the highest percentage of grouping for each behavioural variable. For example, for the “Computer/Internet Experience” behavioural variable, the values with the highest percentage grouping are “Seldom” and “Often”. Having identified these values for all the behavioural variables, we focus on observing which individual subjects most often select the chosen values.

As Fig. 9 shows, the persona Sandra Wolf has been created based on the groupings (illustrated by a solid line) of subjects 1, 4 and 5, because these subjects most often selected the behavioural variable values identified with the help of the percentage grouping table. Using ellipses (solid and dashed lines) to identify the created personas provided a record of the behavioural variable values of which they are composed. This is useful for preparing Sections 5 (Skills & Knowledge) and 6 (Context/Environment) of the Personas Foundation Document and the Narrative as a whole.

In Activity 5 – Synthesize Characteristics and Relevant Goals, we synthesize the details of the data gathered from the interviews designed and run in Activity 1 and from the synthesis of responses to the interviews held in Activity 2 for each of the significant behaviour patterns identified earlier. This synthesis is a specification of brief
points that describe the features of the observed behaviours. This synthesis is set out in the Personas Foundation Document. Having a Personas Foundation document was helpful for classifying the synthesized information about the persona, creating the narrative and completing the annotated use case diagram. Fig. 10 shows the Personas Foundation Document that is output by this activity.
In Activity 6 – Check for Redundancy and Completeness, we check the mappings, characteristics and objectives of the personas to find out if there is any important gap that needs to be filled. To do this, we check that the important identified aspects are fully defined in the personas created and models built through participatory inspection meetings. After running this check on the created personas (Sandra Wolf and John Smith), no information was found to be missing.

After this, we then write a third-person narrative in Activity 7 – Expand the Description of Attributes and Behaviours. The narrative is a clearer way of conveying or communicating people’s attitudes, needs and problems to development team members. Many aspects in the Personas Foundation Document are vague, and it is not as useful for this task as the Narrative. The result of this activity is a Narrative for each of the created personas. The details should expand the list of features in the Personas Foundation Document.
with additional data derived from observations and interviews. The narrative should state what the persona expects from the software system under development and should contain a photograph of the persona that we are describing. Fig. 11 illustrates the Narrative for the persona Sandra Wolf.

In Activity 8 – Designate Persona Types, we select representative personas for requirements elicitation. Based on the description of each of the personas types and all the analyses conducted throughout the personas creation process, we determine the person types (primary, secondary). We associate a type with each created persona: Sandra Wolf is a typical primary persona and John Smith is a typical secondary persona. Then we determine which functionalities raised by the secondary persona and not expounded by the primary persona enriched the software requirements specification. To do this, we analysed the Personas Foundation Document and Narrative for the secondary persona. In this case, no changes were made to the software requirements specification.

In Activity 9 – Build Use Cases, we built an Annotated Use Case Diagram and a Use Case Specification based on the Personas Foundation Document, Narratives and the knowledge acquired from the user throughout all the previous activities. The Annotated Use Case Diagram is based on the traditional use case diagram, to which we have added a brief description of each persona involved in the use case. This description contains the following information for each persona: name and photograph, persona type, sex, age and occupation, skills and knowledge, and goals. We gather this information from the Personas Foundation Document (especially, Persona Identification, Objectives and Skills & Knowledge) and complete the Annotated Use Case Diagram with an informative note on the use case. Fig. 12 illustrates the Annotated Use Case Diagram for the Web Flight Booking System.

We specify the use cases after having completed the Annotated Use Case Diagram, taking into account the particularities and features of the created personas for this purpose. This information is centralized in the Personas Foundation Document. Each specification should have a heading: the name of the persona for whom it has been created. Fig. 13 illustrates a fragment of the Flight Booking Use Case for Sandra Wolf.

Finally, the web flight booking system prototypes are implemented and validated in Activity 10 – Implement and Evaluate Prototypes. We have used the use cases developed in Activity 9 to implement the prototypes. Fig. 14 shows the flight data selection screen for the web flight booking system.

We evaluate the implemented prototypes at participatory meetings. This evaluation was conducted by the future user Elizabeth Brown. Fig. 15 illustrates the mock-up evaluation document for the flight search functionality. This document has several parts. Part 1 describes the tasks to be evaluated. Part 2 contains the user name, day, time and place of the evaluation. Part 3 contains the script for each task recording the order of the actions that the user is to perform and important aspects to be taken into account during the evaluation. The scripts in Part 3 are drafted based on use cases. Finally, Part 4 sets out the conclusions of the evaluation for each task, outlining the changes to be made to the mock-up in order to satisfy user needs.

Personas are only a side topic of requirements engineering, which is again only one part of a software project. Building personas into the requirements stage involves increasing the requirements phase activity workload in order to acquire and use other additional information in return: knowledge and modelling of the user. We cannot expect to add information to the requirements phase activities without increasing workload.

The improvement of software system usability entails increasing the process workload, as we are adding something new (improved knowledge of user features by learning their behaviour patterns). The new information we add is not in place of something else that we can delete. In other words, adding HCI techniques to the development process appears to lead inexorably to a longer process, which will require a heavier workload. In exchange, it will be possible to adapt the software system to aspects of user behaviour and achieve better software system acceptance by potential users. This also applies to our proposal.

Many products, some of which are somewhat time-consuming, have to be output to apply our Personas technique proposal. If time is short or it is too costly to build all the products, it may be feasible to lighten our proposal. The most resource-consuming part of the Personas technique is to hold, transcribe and process user interviews. They cannot be omitted in order to lighten the technique, as it relies on this information. The only way of lightening the interview workload somewhat would be to combine this activity
with interviewing as part of requirements phase activities. Even so, we will be eliciting some knowledge that is not usually elicited in standard requirements-oriented interviews. In any case, we will have to allow extra elicitation time to get to know users, which is the Personas leitmotiv. Obviously, a light version of our proposal will yield fewer benefits. We might omit:

- Percentage Groupings Table, which is the output of Activity 4 (Identify Significant Behavioural Variables), because it is possible, albeit harder, to identify behaviour patterns by just observing the mapping.
- Mock-Ups and Mock-Up Evaluation Document, which is the output of Activity 10 (Implement and Evaluate Prototypes). Mock-Ups can be left out because they are built as part of the requirements phase activities. The Evaluation Document can be omitted because users give useful feedback for evaluation purposes during the mock-up demo, although the information is not as precise as when users perform scripted tasks.

We built a tool to help software engineers use the proposed version of Personas. The tool is publicly available at the web site http://arantxa.ii.uam.es/~sacuna/PersonaSE/tool/. UML has been used to define use cases, class diagrams and sequence diagrams. Some of the use cases that make up the Technique Execution Subsystem are:

- Enter Personas hypothesis.
- Modify Personas hypothesis.
- Interview.
- Modify interview.
- Synthesize interviews.
- Modify interviews synthesis.
- Enter list of behavioural variables.
- Modify list of behavioural variables.

This tool has been developed using Java. The Eclipse and NetBeans integrated development environments (IDEs) were also used. The tool proved very helpful for generating each technique output for the application example discussed in this section. The tool provides forms for each activity and provides the software engineer with guidance for filling each document. Using the tool software engineers can focus on the content of the product of the activity rather than having to remember what information the product should contain.
7. Conclusions

Our research goal was to incorporate HCI knowledge into routine SE practice. Personas is a HCI technique for developing usable interactive applications. We think it is worthwhile to adapt Personas for integration into SE development process. The integration of Personas into the SE requirements stage might improve the understanding of what the software product should do and how it should function.
behave. Personas helps to gain a deeper understanding of the users that are to interact with the system and provides support for developing a system that conforms to user characteristics.

The Personas technique analyses and synthesizes the information related to software system users and models the user. This appears to help focus the software analysis and design activities on the characteristics and objectives of the end user. Personas also appears to be helpful for transmitting key information about users to all the development project stakeholders.

Almost all the Personas steps have a series of shortcomings concerning procedure definition and product description. We have modified the HCI Personas technique to comply with the levels of systematization required by SE.

We have enriched the SE requirements process by incorporating Personas activities into requirements activities. Requirements elicitation and requirements analysis are the RE activities most affected by incorporating Personas.

Acknowledgements

This work has been funded by the Spanish Ministry of Science and Innovation as part of the “Tecnologías para la Replicación y Síntesis de Experimentos en IS” (TIN2011-23216) and “Go Lite” (TIN2011-24139) projects.

References