

Developing and Validating Personas in e-Commerce: A Heuristic Approach

Volker Thoma¹ and Bryn Williams²

¹University of East London, Romford Road, London E15 4LZ

²Betfair Limited, Winslow Road, London W6 9HP

Abstract. A multi-method persona development process in a large e-commerce business is described. Personas are fictional representations of customers that describe typical user attributes to facilitate a user-centered approach in interaction design. In the current project persona attributes were derived from various data sources, such as stakeholder interviews, user tests and interviews, data mining, customer surveys, and ethnographic (direct observation, diary studies) research. The heuristic approach of using these data sources conjointly allowed for an early validation of relevant persona dimensions.

1 Introduction

One of the greatest challenges in user centered design is the elicitation and communication of user requirements. Results from traditional methods such as questionnaires and interviews are often too complex or too cumbersome to digest [1]. Furthermore, design specifications can fairly quickly become out-of-date and may simply not be adhered to by the development team. To resolve these common problems, Cooper [2] proposed the use of “personas” to achieve a more effective tool to elicit and convey user requirements. Personas are fictitious, yet specific and concrete representations of target users. The description of a persona usually consists of a name, basic demographics, motivations, product usage patterns, personal likes, etc. The aim of personas is to inform design and facilitate communication among stakeholders [1]. The persona method has gained a substantial following among developers [3] and is now a standard component of user experience projects in many companies. Advocates claim that personas engage teams to think about users, are easy to use for designers and developers, and help overcome the typical problems associated with more complex or statistical representations of the user base.

The general approach to persona development is to gather information about users’ needs, behaviors, and preferences, which are collated and used to construct vivid descriptions about admittedly fictional individuals [4, 5]. The exact method of how to derive and construct personas is a topic of debate [1]. Grudin and Pruitt [4] advocate a multiple-methods approach, with the use of a range of qualitative and quantitative data to achieve greater accuracy and utility for persona creation. Unfortunately, because of lack of time or resources some practitioners select either purely qualitative or purely quantitative methods in the development of personas. Such limitations have their own problems, such as running the risk of being dismissed by stakeholders as

purely creative exercises [6] or not having the depth of understanding of the users behaviours to generate fully-rounded personas [4]. Finally, it is very difficult to verify that personas are accurate [3]: how can we test whether the combination of attributes actually represents any users? In summary, any persona development project faces three major questions: Which and how many methods to use? Which procedure to follow? And can one validate personas? The main aim of the paper is to illustrate – using a case study – how data gathering methods can be prioritized, adapted and used conjointly for persona development projects. The issue of cost is beyond the scope of this paper, but our results help to focus on core methods for persona projects. A particular focus is on different ways to validate personas early in the research process.

2. Persona Research, Validation and Creation

A large exchange betting company requested a set of company wide personas which would guide design processes as well as product development decisions. Betting exchanges offer online betting in a “many-to-many” form, where users are betting against each other using the betting exchange platform. Additionally, the betting exchange website offers other online products (e.g., poker, casino games).

The process for developing personas consisted of four phases which partly overlapped. The first phase comprised of gathering qualitative and quantitative data and included a review of in-house knowledge. Qualitative data such as observations and interviews were analyzed using a simplified thematic analysis approach [7]. Quantitative data were analyzed using cluster analysis techniques and principal components analysis [8]. The data gathering methods used are summarized in table 1. Phase 2 involved the organization and coding of datasets to allow for the extraction of high-level patterns and themes while preserving the details of the data for later analysis. In phase 3 these themes and high-level patterns were used to define the major dimensions representing our users. Finally, in phase 4 (not described here), the personas were developed in detail drawing on the data and themes extracted earlier.

Stakeholder interviews (semi-structured) were conducted early on in the project. A common theme from stakeholders was that high volume users should be differentiated according to which products they use and their sophistication. In addition to interviews, workshops were held at various stages throughout the persona development process. From the results of these activities and other previous requirements engineering projects (customer surveys, market segmentation) we compiled a long list of categories and parameters that would potentially be used to describe personas, such as interests, betting behaviour, psychographics, motivation, media profile, and many more. Over 200 categories and subcategories were recorded and were entered in a dynamic “master” spreadsheet and regularly updated.

We then carried out user interviews (n=24) and also directly observed eight users in their natural context to elicit typical motivations, needs, context of use, etc. Home visits were arranged in advance and users were instructed to keep a week-long “diary” of their betting related experiences. These methods provided direct context-driven information that would be difficult to otherwise uncover (e.g. sharing of accounts, social aspects of betting, etc.). The extracted themes were around the expertise with betting and emotional involvement with a particular sport or product. In addition, a

clear distinction was found between sport-only bettors and those also using the casino/games products. A further method to elicit persona information was user testing. Although usually centered on prototype development and testing, we designed the user tests to probe clients for persona relevant information. Emerging themes here were the reliance on old navigation habits, the lack of awareness of other products, and betting related features on the site that made betting more complex. User tests also validated themes that had previously emerged. For example, the inclusion of minimized games within the sports betting page was tested on users with different product usage profiles. Five out of six casino users spotted the minimized game, whereas only one out of six non-casino users found the game, thus validating the dimension of product-awareness and usage. A further way to validate previous themes was through data mining. An analysis of navigation paths showed that the way to navigate to customer's favorite markets depended on how experienced they were and how often they betted, validating themes extracted from user tests. Finally, a customer panel with over 1000 users provided feedback to electronic and phone surveys. This data base was linked to the data warehouse which allowed for the analysis of usage patterns across clients. Quantitative patterns (clusters) emerged on volume of product use (e.g., sports betting, games, casino, etc.) and channels (online, mobile).

After data were collated from all sources, emerging themes and patterns were extracted [7] and summarized as various persona "dimensions" (see [1,5] for structuring persona attributes). Initially 10 dimensions were extracted and then reduced to four, which were mapped onto a "persona space" (Figure 1). A persona could vary in: 1. betting volume, 2. passion or emotional involvement in a product (e.g. horse betting), 3. degree of betting sophistication (from punting to complex and automated trading), and 4. type of product usage (e.g. sports betting vs. Casino).

Fig. 1. Persona space

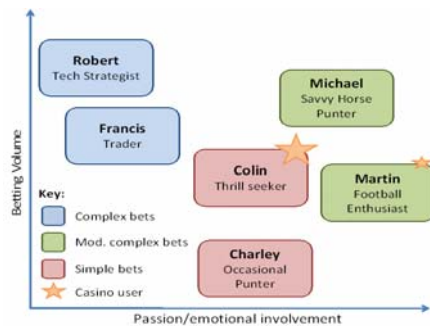


Table 1. Methods used in persona research

Data (phase)	Recommend for
Stakeholder Interview (early)	Themes
User tests/interviews (throughout)	Qual. & quant. data, validation
User observations User diaries (early)	Qualitative data, themes, quotes
Stakeholder workshop (early/late)	Qualitative data, validation
Data mining (throughout)	Quant. data, validation
Customer survey (throughout, late)	Quant. data, validation

To summatively validate these dimension, 24 variables from the user panel response (n=408) were subjected to an exploratory factor analysis (oblique rotation). The variables represented quantitative information on usage patterns and Likert-scaled answers on betting related questions. The first of the four main factors extracted loaded highly (partial correlations > .48) on variables such as: Joy from betting, strategic betting, betting as income, information based betting, betting on specialty areas. The second factor loaded (> .20) on: Revenue (volume), profit & loss, mobile betting volume, in-play betting volume, Casino betting volume; The third factor loaded on (> .16): Age, Internet experience, technical know-how, betting

experience, frequency of Internet use, betting habits, product loyalty. The fourth factor loaded ($>.17$) on: Perceived lack of know-how, betting as entertainment, amount of online gambling; and games volume. This result confirmed the previously derived themes from qualitative sources (the variables loading on the fourth factor reflect the various motivations of using different products such as games and casino). The dimensions and the “locations” of different personas were further validated and fine-tuned in stakeholder workshops. The persona space greatly facilitated the validation and communication of persona dimensions to stakeholders. In addition, it allowed the team to quickly produce and test early persona sketches.

3 Conclusions

We began this article by describing the three main problems of persona creation: the choice of methods, the procedure, and validation. We propose a heuristic approach that includes a limited number of both qualitative as well as quantitative methods, including user observations, user tests, in-depth interviews, data mining, and surveys. Persona development must be based on both qualitative and quantitative data if a fully rounded and believable set of personas is to be developed [1, 4]. A second important point is rather than using a particular method in isolation, we recommend using qualitative and quantitative data gathering methods in parallel to extract relevant themes. Further, we propose validating the extracted persona dimensions early on in the process: through triangulation (comparing with themes from different methods) as well as by using quantitative data reduction (here: factor analysis). The use of a persona space to represent relevant dimensions is a pragmatic solution to validate that personas reflect true user groups, while at the same time providing a framework for how to conduct research and integrate the findings across methods [4]. We contend that although we may not have validated personas in a fully scientific sense [3] or exhausted other (qualitative) approaches of validation, the approach described here is a practicable approximation to ensure the creation of relevant and cogent personas.

References

1. Pruitt, J., & Adlin, T.: *The persona lifecycle*. Morgan Kaufmann, San Francisco (2006)
2. Cooper, A.: *The inmates are running the asylum*. Macmillan, New York (1999)
3. Chapman, C.N; Milham, R.: *The personas' new clothes*. Human Factors and Ergonomics Society (HFES) 2006, San Francisco, CA. (2006).
4. Grudin, J. & Pruitt, J.: Personas, participatory design, and product development: An infrastructure for engagement. *Proc. PDC 2002*, 144--161 (2002)
5. Pruitt, J., & Grudin, J.: Personas, Practice and Theory. *Proc. DUX* (2003)
6. Mulder, S.: *The User Is Always Right: A Practical Guide to Creating and Using personas for the Web*. New Riders, Berkeley, CA (2007)
7. Braun, V., & Clarke, V.: Using thematic analysis in Psychology. *Qualit. Res. Psy.*; 3, 77—101. (2006).
8. Hair, J.F., Anderson, R.E., Tatham, R.L., Black, W.C.: *Multivariate data analysis*. Prentice Hall, New Jersey (1998)