

Global user research methods

Tjeerd de Boer

**Additional Contributors for this chapter are listed with their corresponding sections.*

6.1 INTRODUCTION

This book is primarily concerned with moderated usability tests because this is probably the most popular global user research method. However, there are numerous other methods that can be used for global user research as well. There are interesting aspects to each of these methods, but a complete overview of all methods is beyond the scope of this chapter, so the focus is on those that are more frequently used.

An overview of the global user research methods can be given by mapping the global user research methods on several key dimensions (Rohrer, 2008). For the purposes of this chapter, we have used two of the dimensions that were put forward by Rohrer:

Behavior versus attitude

Qualitative versus quantitative

The distinction between behavior and attitude can be translated to what people do versus what people say, which can be very different. Usability research is typically on the behavioral end of the spectrum, but it can still be useful to have information on what people say. For example, online surveys can help you identify how users rate your product or service against competitors.

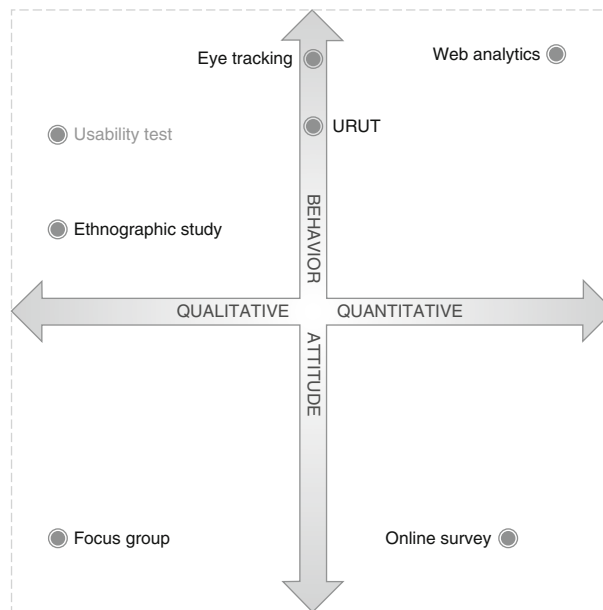
When doing qualitative research, you gather data *directly*. For example, during ethnographic research you directly observe the behavior of people. The analysis of the data is usually not mathematical, but it does give you insight into the reasons for observed behavior or

possible solutions for problems identified. In contrast, quantitative research typically uses a tool for data collection (e.g., Web analytics and online survey). This *indirect* research gives large amounts of numerical data, which allows you to quantify behavior and to perform quantitative analysis on the data.

Figure 6.1 illustrates where the global user research methods appear along these two dimensions. It only shows the research methods that are addressed in this chapter, but other research methods can be placed along these dimensions as well. The usability test is shown as a reference point.

Each method in this chapter is discussed according to the behavior-attitude and qualitative-quantitative dimensions. This gives the following structure for the chapter:

- Ethnographic studies
- Focus groups
- Eye tracking
- Unmoderated remote usability testing
- Web analytics
- Online surveys



■ **FIGURE 6.1** Mapping of user research methods on two dimensions.

The chapter concludes with a section on personas. Strictly speaking, personas are not a research method. Personas drive design. However, because personas are often used for global user research, we thought a treatment of personas in a global context would be both interesting and beneficial.

6.2 ETHNOGRAPHIC STUDIES

Written by Mercedes Sanchez

6.2.1 Introduction

Although ethnography has its roots in anthropology, the idea of the ethnographer huddled among the native population on a tiny, far-off island is somewhat dated. In today's business world, ethnography has taken on increasing importance as a research method that produces a detailed picture of how products are actually used by people and shows ways in which they can be improved to better meet consumer needs and wants. Unlike focus groups and other qualitative research methods, ethnography takes place in the natural context of the users, which allows you to directly observe what people actually do and not what they say they do.

The goal of an ethnographic study is to directly observe and collect the stories people tell to help them make sense of the world around them. These stories are then given color and insight based on your detailed field observations. The focus of ethnographic study as it pertains to user experience research is often directed to investigations of how specific products or types of products are used. Ultimately, these investigations aim to identify possible product innovations or discover potential market niches. As such, ethnography is being used increasingly more often by companies to inform design and brand decisions and to design products that fill the gaps between those needs a product currently serves and those that users want served, preferably ahead of the competition.

Ethnographic research is well suited for finding issues that other research methods cannot because it is not as focused as most other research methods (e.g., surveys, usability testing, and focus groups). This can be particularly useful when conducting global user research because you often do not know what information you are looking for. The following example about a leading Brazilian home appliance manufacturer illustrates this point. In recent years, this company had set a goal of exporting refrigerators to India. It was through observation of real local Indian consumers that the company learned that refrigerators for the Indian market must have a lock and key because

of a peculiar local situation: Monkeys enter the houses, open the refrigerators, and escape taking some food with them. Although this issue might be straightforward for people living in some areas of India, this insight was definitely new for the Brazilian company. This issue could probably not have been identified by conducting an online survey or a usability test.

Ethnographic research can also be used to explain differences between countries. The same Brazilian company had been exporting stoves to many countries, and its Brazilian models were the only ones that had covers on them. In an effort to reduce production costs, the company had to decide whether or not to eliminate the covers on all Brazilian stove models. However, after an ethnographic study of real Brazilian consumers, researchers advised the company's production team not to do away with the covers because they are very much used and appreciated by Brazilian housewives. These housewives are very mindful about their homes and especially their kitchens. They enjoy having flowers and an embroidered cloth on the stove cover as a clear display of carefulness, which also gives them a sense of "mission accomplished" after cooking a meal and cleaning.

6.2.2 Overview of the technique

There are a number of ways to practice ethnography, and the most successful studies often consist of some combination. The bedrock of ethnography is the combination of informal and formal observation. Informal observations introduce the least amount of bias from the researcher and are obtained mainly by the ethnographer going into the field without presenting himself or herself as a researcher to simply observe the surroundings. Informal observations contain little to no interaction between the researcher and those who are being observed because the ethnographer strives to blend into the background.

Usually, however, somewhat more formal observations are used to gather ethnographic data. Formal observations often include shadowing informants as they go through the normality of everyday life or, in the context of a specific project, behave as they normally would when interacting with a product. In more formal types of observations, informants know they are being watched by a professional researcher, but the researcher relies on the informant to dictate the direction of the interaction and generally only interrupts the flow of events to ask clarifying questions. By following the informant's lead, the

researcher introduces less bias in an ethnographic study compared to, for example, a usability test.

Ethnographers also rely on interviews with informants, though these interviews tend to be relatively informal and either unstructured or semi-structured. Importantly, ethnographic interviews rely on storytelling by the informant and the ethnographer should attempt to stay out of the conversation as much as possible.

Ethnographic interviews take place in the field, typically occur after an event has taken place so as not to disrupt the flow of the event, are in great depth, and are fairly conversational. Ethnographers will generally go into the field with a list of topics to discuss with informants, but the conversation's impetus should come from the informants. Although ethnographic studies yield the best results when they are longitudinal, this is not always possible with user experience research projects due to time and budget constraints. This does not mean, however, that ethnography cannot be used to great effect on short-term projects. It simply becomes necessary to increase efficiency. To do so, short-term projects often focus on a specific type of activity, such as running errands, shopping online at home, or observing specific types of work behaviors. Prior to the study, participants are notified of the general goals and are scheduled according to when they might naturally be performing the actions to be observed. Focusing participants on specific tasks means the ethnographer will have a better chance of capturing the types of behaviors and stories necessary to make good recommendations.

6.2.3 **Fieldwork**

Observing people is not as simple as it may seem. There are myriad nuances that inform human behavior, and these often rely on culturally specific habits. It is important for you to enter the field with knowledge of the product being studied and the type of informant being observed. Each stage of the study, from screener to analysis, has to have been considered to ensure that relevant data are collected. Most important, you must go into the field with open eyes and an open mind.

The most important component, and often the most difficult to achieve, to gathering rich data from your fieldwork is building a meaningful rapport with your informant. For longitudinal studies, a deep rapport can develop over time. However, time is often the one thing that user experience researchers do not have. As a result,

the initial screening process is critical to recruiting informants who are excited to bring you into their daily lives and eager to share their routines. However, you must realize that by screening for the most eager informants, the opinions of those who are less vocal are lost in the analysis.

Another important rapport-building practice is to make sure the participant is aware of your project and the project's goals before going into the field. It is important that you are open with the informant about what fieldwork will entail, included activities to be covered and how they will be captured, and to get feedback from the informant before meeting with him or her. You should attempt to have conversations with the informant before going into the field so that both the informant and you are more comfortable upon first meeting each other. If possible, you should ride along with the informant to the site of the fieldwork in order to begin building a picture of their habits and behavior.

The most important thing you can do to build rapport with informants is to remind them of the knowledge they possess and that you are there to learn from them. It is intimidating for informants to have researchers come into their lives to observe them. They may be even more intimidated when the ethnographer is holding a camera. To put informants at ease, remind them of their own expertise and that their comments and insights will help you in gathering valuable information. When doing ethnographic research, the informant really is the teacher.

An ethnographic study generally involves more than observing, and researchers should make use of other sources of information. If possible, researchers should plan to do at least one initial meeting, either individually or in groups, and one individual interview with each participant at the end of the study to follow up on any questions or issues that were not discussed during fieldwork. It is usually a good idea to discuss observations or insights with informants to gather more feedback from them.

When collecting data, it is also important to do more than just take notes. For example, ethnographers often observe informants after weeks of writing a journal or keeping a photo or video diary. These artifacts can then be studied by ethnographers before entering the field so that they have some knowledge of the informants. Also, these artifacts are good talking points in the field during lulls in the

conversation. Because you are in the informant's natural environment, it is possible to ask questions directly about elements in that environment. If possible, you should take a video camera into the field to record interactions because handwriting notes inevitably leads to missed data. In the end, these types of materials will be very valuable for analyzing and presenting results.

6.2.4 **Planning a global ethnographic research project**

When conducting global ethnographic research, the basics outlined previously apply. However, when working internationally, it is important to work with local partners. Even if you speak the same language as your informants, important cultural clues can be easily missed due to different cultural backgrounds. Because you are working with partners, it is vital to have a detailed description of the type of informants to present to them. This will help ensure the quality of informants, which will ultimately lead to richer data. Working with a local firm also helps to ensure that specific questions relating to the local research environment are not overlooked. For example, in Brazil it would be easy to obtain permission from participants to enter their homes to study an issue. In other countries, such as The Netherlands, this would not be so easy. Although it is sometimes possible to attend the fieldwork with the local ethnographer, having an additional observer who speaks a foreign language will do little to make the participant feel more at ease. Instead, rely on the insight and knowledge of your partner firm.

A company may decide to do a global study and send its own ethnographer to attend the fieldwork with the local ethnographer. In this situation, a greater chance of success is assured if there is no attempt made to simultaneously translate the interaction between the local ethnographer and the informant. Doing so can distract or intimidate the informant. Even without simultaneous translation, the combined work can be productive because the local ethnographer will observe some behaviors that the foreign ethnographer (you) could not perceive, and the foreign ethnographer may catch cultural insights the local ethnographer takes for granted. Even when you do not speak the local language, you can observe the visual details, the nonverbal clues, and get an interesting picture of the informant's behavior. It is important that you discuss these observations with the local researcher afterwards because nonverbal clues can differ between cultures.

6.2.5 Analysis

Analysis and presentation of ethnographic data are challenging, but the results are usually well worth the effort. Although it is easy to simply present data in descriptive reports, ethnography's real power lies in its ability to go beyond simple descriptions and to find patterns of behavior. Ethnographic reports should highlight gaps both between the way companies believe their technology is used and the way in which it is actually used and between the way technology works and the way informants actually want or need it to work. Based on these observations, the report can make suggestions on how to fill those gaps. Ethnographic results can be presented in a number of ways, including posters, videos, photo collages, storyboards, user profiles, and personas. These means of data presentation are well suited for ethnographic research because they do what good informants are supposed to do – tell meaningful stories about real situations.

6.2.6 Conclusion

Ethnographic research is well suited for identifying issues that other research methods cannot. By entering the field with an open mind and open research protocols, you can gain insights you had never before considered. Ethnographic research is particularly useful when conducting global research because you often do not know what kind of issues could be found in countries with which you are unfamiliar.

When conducting ethnographic research outside your own country, it is recommended to use local partners in the study. This not only solves language issues but also ensures that subtle cultural clues will be noticed.

6.3 FOCUS GROUPS

Written by Mercedes Sanchez

6.3.1 Introduction

Focus groups, which basically consist of groups of people brought together to discuss opinions, issues, and concerns about something, are among the most well-known qualitative research methods and have been in use for many years. Although most frequently linked to market research and advertising, focus groups can in fact be a powerful tool to discover user needs, goals, feelings, attitudes, preferences, initial reactions, and desires from specific target audiences. Focus groups can be held before, during, or after the product or site development, but

they are usually conducted at the beginning of a project cycle. In these situations, focus groups are very suitable to help identify and prioritize requirements so that key customer wants get addressed first.

Focus groups are a good technique to obtain user opinions and to discover what they want, but they are a poor method to discover usability problems. This is because group discussions are focused on what the users say, instead of what users do and how they interact with a system or a product. It is therefore recommended that focus groups be used in combination with research methods that focus on the behavior of the users (e.g., usability test, ethnographic research, and Web analytics).

Focus groups can also be used when doing global user research because it is sometimes not sufficient to test an interface for usability problems. You may need a deeper understanding of the local users to compliment the data obtained from the testing (e.g., needs, goals, and feelings). This is when the focus group is very useful because it is much more efficient to have a focus group discuss the problems in a single session than to organize field trips to the locations of the users.

An alternative to conducting focus groups is to organize online group discussions. This can be more practical when the users are located in different countries, but it has a number of disadvantages. Online discussions are discussed later.

6.3.2 Overview of the technique

In a typical focus group, 5–10 participants discuss their thoughts, preferences, and concerns about a product or an interface. Each group typically lasts 2 hours and is conducted by a moderator, preferably local, experienced in keeping discussions on track, and with good knowledge of the subject in question. Normally, four to six groups are necessary to get representative results and to begin to see convergence in the qualitative findings.

6.3.3 Considerations for conducting global focus groups

Defining research objectives

When defining research objectives, keep your list short. Avoid including too many objectives or subjects just because it is an opportunity to “talk with your customers.”

Objectives must be clear and focused on what you want to discover. This is particularly important in global studies because there may be different objectives for each country. The larger the list of objectives, the smaller the chances you can accomplish what you need.

Planning research

In a typical local research project, four to six groups are necessary to get representative results. When dealing with global user experience studies, one must plan at least three homogeneous groups per country to get representative results in each country.

Recruitment

The most important issue regarding recruitment for focus groups is to define the user profile of each group. The user profile (or profiles) must be representative of the target audience.

Except in special cases (e.g., when you want to investigate gender or age conflicts), groups are usually composed of homogeneous participants in order to maximize disclosure.

In different countries, local culture must be taken into consideration when defining the profile of groups. Thus, when establishing selection criteria for global groups, consider aspects such as the following:

Gender: Will men and women feel comfortable discussing the subject in a mixed-gender group? For example, this could be an issue in India and in Arabic countries.

Age: How intimidating would it be for a young person to be included in a group of older participants or vice versa? For example, this could be an issue in Japan and China.

You also have to cover the basic procedures of recruitment for global research, as discussed in Chapter 3.

Moderation

Provide a moderation guide with explicit objectives and goals for each round of the discussion, point out questions that cannot be overlooked, and discuss the desired time to be spent on each round. Discussions must be conducted by local and experienced moderators with the ability to “speak the same language” of the participants in order to elicit spontaneous reactions and ideas from the group.

If the object of the discussion is a Web site, a mobile phone, or a system, be sure that the moderator is familiar with the subject. Moderators

used in market research may not have sufficient knowledge to conduct discussion on technological issues or user-centered design.

Consider issues about gender and age when defining or contracting the local moderator. For instance, in some countries men are not used to accepting orientation or questions from women, so the moderator of a group composed mostly of men should not be a woman in this case. In other countries, young people do not talk before older people do, so a young moderator should not conduct a group of older participants.

Chapter 4 provides useful information about moderation of global studies.

Translation and recording of sessions

The screener and moderation guide must be translated, and the translations must be checked. In countries with different languages from yours, you may need to have a simultaneous translator for sessions even if you are not attending the discussions, in case you need the audio recordings in your language. If you do not need the audio recordings in your language, you can substitute the (costly) services of simultaneous translation for a skilled note taker (less costly than a translator) fluent in both languages.

Usually, discussion sessions are recorded in video and audio and then transcribed, which is expensive and takes quite a long time. When attending sessions in many countries, this can compromise your deadline for results. An alternative is to have a member of your team take copious notes during sessions or use a local qualified note taker who can write in your language.

Analysis

When doing many sessions in different countries, you may obtain a huge amount of information, so it is prudent to create a unique template for reporting the discussion results by all countries involved. With the results of all countries in your hands in the same format, the analysis is easier and takes less time to perform.

6.3.4 Focus groups and usability testing

Group discussions are focused on what users say, and usability testing is focused on what users do – how they interact with a system or a product. Focus groups are therefore a good technique to use for obtaining users' opinions and to discover what they want, but they

are a poor method of discovering usability problems. If your list of research objectives has items such as “discover how users enjoy the new feature” and “discover if users can use the new feature easily,” you will have to combine focus groups, to gather reactions, and usability testing, to observe how users actually deal with the new feature.

You must consider when focus groups will be the most effective during a project. In general, focus groups are done prior to usability testing, but they can also be used together, for example, during a redesign phase. After discovering what works and what does not work on an interface by observing individual users, a group discussion can provide insights to innovation and a deeper understanding of the needs and desires of consumers from different countries. For more complex systems, this may be the best approach because it can be easier for users to give feedback on specific items after having tried the system.

6.3.5 Alternatives to focus groups

An alternative way to have group discussions with the users is to organize online discussions. This can be more practical when the users are in different countries because you do not have to travel to all these locations.

Online group discussions can be done by posting questions to a newsgroup, online forum, or an e-mail discussion list. A disadvantage of this approach is that these online discussions are publicly accessible, so they cannot be used for confidential information. Another problem is that users active in newsgroups, forums, and e-mail discussion lists might not be a good representation of your users.

A solution to these problems is to organize group discussions in an online panel. An online panel is different from newsgroups, forums, and e-mail discussion lists because it is not public. People can only participate in the discussion if they are given access to the secure environment. The participants can be recruited to be representative of your users, so there is no bias toward very experienced Internet users.

A major disadvantage of all forms of online discussion is that you can only see the explicit feedback of the participants (e.g., usually written comments). Using focus groups, you can observe the nonverbal behavior as well. Using video conferencing software might eliminate this disadvantage, although the quality of inexpensive video conferencing software is usually not adequate. This situation will probably

change in the near future as inexpensive high-quality software and the necessary broadband Internet connections become widely available in most areas of the world.

6.3.6 Conclusion

Focus groups can be used when doing global user research because it is sometimes not sufficient to test an interface for usability problems only. You may also need a deeper understanding of the local users to compliment the data obtained from the testing (e.g., needs, goals, and feelings). This is when the focus group is very useful because it is much more efficient to have a focus group discuss the problems in a single session than to organize field trips to the locations of the users. An alternative to conducting focus groups is to organize online group discussions or online panels.

6.4 EYE TRACKING

Written by Aga Bojko

6.4.1 Introduction

The popularity of eye tracking as a user research tool continues to increase. Once thought of as a nice-to-have, eye tracking is slowly turning into a must-have, typically employed in combination with other methods or, less often, as the only data collection method in a study. Eye tracking in user research is used to capture the eye positions of a person looking at a visual stimulus, such as a Web site, an image, a video, a product, or even physical surroundings. Based on these eye positions, we can infer where the attention was directed and which information was being processed. By examining the spatial and temporal relationships between the eye positions, we can also understand *how* the information was being processed (e.g., whether a section of text was skimmed or thoroughly read, or in what order the display elements were noticed).

In Figure 6.2, the participant (right) is sitting in front of the Tobii 1750 remote eye tracking system integrated into a computer monitor. The session moderator (left) is sitting in front of the control station that enables her to view the participant's eye movements and control the eye tracking software. The blue dot on the moderator's screen indicates the point of gaze of the participant.



■ FIGURE 6.2 Eye tracking session in progress.

Eye movement data are especially valuable for answering questions for which conventional measures of user experience are not sensitive or accurate enough. One of the main reasons to track eye movements is to understand the user experience that takes place during and between behavioral events that are observable without the help of any tools. These events include mouse movements and clicks, physical object manipulations, and body movements. Although researchers can try to describe users' cognitive processes affecting the user experience based on the information verbally offered by the users, eye movements provide data that are much more detailed and objective than user self-report. Eye tracking can help us understand user expectations, interest, confusion, and decision-making processes, which provide insight into the effectiveness of a design.

Equipment, skills, and knowledge pose many challenges in eye tracking studies. These challenges are magnified in studies conducted across multiple geographic locations. First, eye tracking requires specialized equipment that most research labs do not yet have. However, even if a lab has eye tracking equipment, it may be different from the equipment that is needed for the research. For example,

an eye tracker built into a computer monitor is not appropriate for tracking user interaction with the surroundings (e.g., when shopping, driving, or playing sports). Instead, a wearable eye tracker should be used, and if it is unavailable, the study simply cannot be conducted.

The second main challenge in eye tracking studies involves the researchers, who need to have the appropriate skills to operate the eye tracker and collect data, as well as the knowledge to plan an eye tracking study (e.g., select the correct procedure and measures) and analyze and interpret the results. There are several commercial eye trackers on the market that are fairly easy to use. However, the fact that someone can “track participants” and “generate heat maps” does not mean that he or she has any understanding of the nature of eye movements. In this section, we offer advice on how to overcome the equipment- and skill-related challenges in global-scale research.

6.4.2 Equipment-related challenges

The lead team may either decide to bring its own eye tracking equipment to the other study locations or use equipment provided by the local teams. Both solutions have a number of advantages and disadvantages, and they require careful preparation to ensure a successful study.

When lead team brings own eye tracking equipment

When taking its own equipment to another country, the lead team should be aware that some of the equipment may not successfully transfer through customs. It may also become damaged or even lost on the way. Although one can conduct a usability test and collect valid data without an actual usability lab, an eye tracking study cannot take place without eye tracking equipment. Sometimes, all that the lead team can do is plan to arrive early and hope for the best. Setting the right expectations with the study sponsors and the team will minimize surprises and ensure that everyone is aware of the risks involved in traveling with the equipment.

Nevertheless, it is always advisable to have a backup plan in case the equipment does not arrive with the lead team. If the local team does not have the proper equipment, one option is to find a local company that has a similar eye tracker to the one used in the study. Sometimes equipment can be rented at the last minute if the lead team or the local team communicates with the company ahead of time and explains the plan. The rental may not be the exact hardware model or software version that is needed, or it may be expensive. However,

this can mean a difference between being able to collect data or not, which is especially significant for studies with eye tracking as their main or only component. Setting plan B into motion will likely require time for transport of the backup eye tracker, setup (which could differ from what the lead team is used to), and possibly some self-training, so the researchers should plan accordingly.

Assuming that the lead team's equipment made it to the destination unharmed, it is now important to remember that the hardware may have a power supply incompatible with the power supply used in the test countries. For example, in Europe and mainland Asia, the supply voltage ranges from 210 to 240 V and can damage equipment set for North American or Japanese voltage of 100–120 V. If the equipment has a voltage switch, the switch should be flipped to the correct setting before plugging in the system. If there is no such switch, a voltage converter is needed. A simple adapter plug will not solve this problem.

If the study requires participants to interact with a Web site, software application, or images on a computer, and the lead team is bringing its own test computer (which may be wise if the computer is also running the eye tracking software), the operating system (OS) and keyboard could be different from the OS and keyboard used at the destination. Asking participants to type on a differently configured keyboard or use an OS in another language (even one with which they are familiar) may produce artifacts in the results, depending on the study focus and methodology. Although these artifacts may be of little significance in a qualitative usability study, their impact on an eye tracking study will be far greater.

For example, imagine that every time a participant tries to type a "Y," a "Z" appears instead, or several letters of the alphabet are entirely missing from the keyboard. As a result, he or she will spend much more time looking at the keys than usual. Every time users look down and then back up at the screen, they need to visually reorient themselves, which requires additional eye movements and produces different scan patterns. If the lead team suspects that using a different keyboard and OS may affect the results of the study, a local keyboard should be used instead and an OS in the local language should be installed.

Sometimes, the lead team may decide to bring the eye tracker only and rely on the local researchers to provide additional hardware (e.g., PC to run the eye tracker). This is a risky option because no

one can check whether everything works together until the eye tracker arrives at the test location prior to the study. Therefore, the lead team should be extremely specific regarding the parameters of the required hardware to ensure that it is compatible with the eye tracker.

When local eye tracking equipment is used

If the eye tracking equipment is available on-site, using it to conduct the study should be strongly considered. However, the lead team must diligently check and confirm that the correct equipment is indeed available. To be able to combine the eye movement data across several test locations so they can be analyzed together, all data may need to be collected using similar (or sometimes even identical) hardware and software, including the software version and settings. The lead team should not make any assumptions but, rather, clearly communicate its requirements to the local teams and make sure these requirements can be met.

If local equipment is used for the study, the lead team needs to decide how it will take the data off of the local equipment. Eye tracking data files can be massive in size – far too big to fit on a DVD or be easily uploaded or downloaded. The local research teams can be provided with portable hard drives that can be shipped to the lead team after data collection is complete. Shipping of the data can have an impact on the schedule, especially if they are shipped across borders. The project timeline and team expectations have to be properly set.

6.4.3 Skill- and knowledge-related challenges

Many companies, even those that have eye tracking labs, might not have adequate knowledge or training necessary to set up and conduct a particular study. The lead team should provide the local teams with very clear instructions. The schedule should also allow for time to train the local team not only how to run the study but also how to do basic equipment troubleshooting.

In addition, the lead team should provide a detailed description of the procedure so that data are collected in a consistent way across locations. The description should include information such as when to start and stop the eye tracker, whether or not to use a think-aloud protocol while eye tracking, and whether or not the moderator should assist the participant if he or she is struggling with the tasks. If left to the local moderators, these decisions can have a detrimental impact on the eye movement data.

Also, any inconsistencies will produce unnecessary variability in the results, thus making them difficult to analyze and interpret.

6.4.4 Conclusion

As the popularity of eye tracking rises, the challenges associated with conducting global eye tracking studies will decrease – the appropriate equipment will become more available and the skills and knowledge will become more widespread among user research professionals. However, nothing can replace careful planning and attention to detail, which will always be necessary to make eye tracking research, especially on a global scale, successful.

6.5 UNMODERATED REMOTE USABILITY TESTING

Written by Jayson M. Webb and Alfonso de la Nuez

6.5.1 Introduction

Today's Web business is far more advanced and sophisticated than it was during the dot-com boom years in the late 1990s and early 2000s. As such, Web managers and marketers have demanded new and more powerful tools to manage their customers' interactive online experiences. Unmoderated remote usability testing (URUT) is an approach born precisely of that demand, and it addresses the following issues that online user researchers might have:

- "I'd like to obtain statistically significant data to quantify usability and user experience. Traditional usability testing in a lab has value in some situations, but the sample is small and it only allows for usability problem identification, not quantification."
- "I'd like users to participate in their natural context, at home, using their own PCs."
- "I have a global audience and would like to test in several countries, but I don't have the budget or the time to conduct a traditional moderated study in a lab."
- "One of the challenges for us now is to be able to do benchmarking. We'd like to compare different competing sites and see who is performing best, offering the best usability and user experience."
- "Web analytics tools give us lots of data about what happened on our site at an aggregate level, but we still don't know why it happened or how it relates to different user goals. We want to know users' real goals and why they do what they do."

URUT allows companies to test Web sites with a large number of geographically distributed participants in their natural context. URUT technology automatically presents tasks to participants and tracks their interaction with a site, including navigation path, page scrolling, and click location. The data collected measure the important usability dimensions of effectiveness, efficiency, and subjective satisfaction (Bevan, 2008; National Institute of Standards and Technology, 2007).

URUT has the following characteristics:

- It is a research solution that uses technology to allow researchers to manage online (remote) usability studies.
- A special kind of software asks users to complete tasks and asks specific questions related to the task (e.g., survey) and the usability of the site. The software supports counterbalancing and randomization of task presentation orders and question orders.
- The same software collects session data as users interact with the site (e.g., Web analytics).
- It allows testing over a large user base, geographically dispersed, so it is ideal for both nationwide and global testing.
- No moderation is required.
- Users can participate at any time of day and in their natural context, using their own PCs.
- Because of the large number of participants, URUT is ideal for quantitative evaluation of Web sites, especially comparisons of sites and benchmarking. In addition to the quantitative data, a large amount of qualitative data can also be collected.
- The rich data provide statistically powerful results on which to base business decisions.
- With the help of online tools, researchers can analyze the data and create a report.

Because users do not need to physically attend the sessions at a lab facility or testing room and there is no moderation needed, cost-efficiency is clearly one of the main advantages of URUTs, but it is not the only one. For researchers and Web marketers, the ability to quantify usability is very valuable. Also, many of today's researchers and online channel managers agree that having users in the lab environment is not always the best solution to obtain quantitative data that generalize to real-world usage. Traditional moderated usability testing is still an indispensable tool and should be part of the

portfolio of user research approaches. However, especially for global testing, URUTs offer research capabilities that were unthinkable a few years ago.

6.5.2 URUT contrasted with other approaches

It is important to understand what URUT is and what it is not in order to use it correctly for user research.

- URUT is not moderated remote usability testing: Many people understand the concept of “remote usability testing” as an alternative to more traditional user testing by adding a Web camera and communication software between the moderator and the end user. Also called “moderated remote usability testing,” these types of studies are similar to traditional testing. The essential difference between traditional usability testing and moderated remote usability testing is the geographic issues resolved and the associated cost reduction. However, it is quite different from URUT, and people should be fully aware of this. Basically, what makes URUT a completely different method is the lack of a session moderator. In moderated remote testing, there is still a need to moderate, so the technique is typically used to test small samples and therefore does not serve to quantify usability.

Moderated remote usability testing can be difficult when doing global testing because one needs a local moderator from each country to carry out the studies, which increases expense. Of course, it is possible to use one moderator for all countries, but this can cause problems with language (e.g., non-native speaker) and the understanding of local issues (see Chapter 7). Using URUT, however, you just need a review of the protocol or pilot testing in the local country before launch, which is a more cost-efficient approach.

- URUT is not an online survey: Online surveys can ask users about goals and perceived success (iPerceptions, 2008) but do not measure actual behavior and actual success. Also, online surveys do not present tasks or follow an entire session and ask questions based on behavior. URUTs, on the other hand, use highly sophisticated software to ask users to complete certain tasks (task-based testing) or accomplish goals (true intent testing and free search testing) and track users, and they can present questions before, during, and after the task or goal has

been accomplished. There is a major difference between filling out a survey at a random point in an interaction with a site and filling out a survey after having interacted with the site in a known way. By using a predefined event to trigger a survey, the survey questions can be more specific and directed. The external validity of the results is thus increased because the researchers know with a higher degree of certainty that participants are responding based on their experience with a specific aspect of the site rather than generalizing across their overall experience.

- URUT is not just another Web analytics tool: Web analytics tools can provide extremely detailed data with respect to user behavior on the Web site. For example, they can measure the number of visits, the location from which users arrived at the site, how much time is spent per visit, and where users abandon a flow. Although it is very important to obtain real-time data regarding what is actually happening on the site, it is also very important to know why it happens. URUT accomplishes this by capturing users' intentions or goals, whether each goal was actually achieved (effectiveness ratios), how much time and effort was dedicated to each goal (efficiency ratios), why users abandoned a specific task, and, finally, how satisfactory the online experience was. By comparing the user behavior data and survey data with the tasks and goals of the users, URUT can be much more than just another Web analytics tool.

6.5.3 URUT in detail

A firm grasp of the practical details of conducting URUT is essential for planning your global user research and prioritizing the tasks for your URUT project. Here, we share our experiences with the effectiveness and cost savings realized by using URUT for global testing and point out some of the problems you might encounter.

Participants and locations

Although a major benefit of URUT is that large numbers of geographically distributed participants can be tested in a cost-efficient way, you still need to understand the local context and constraints in each country before you begin. If you have never done any type of testing in your chosen countries, proceed with caution. If you use a vendor,

make sure the vendor has relevant global testing experience. URUT does not free you from many of the concerns that affect other types of global testing. These issues include translation of testing materials and participant responses, cultural differences in expectations about giving feedback, and expectations regarding written privacy policies (Courage & Baxter, 2005; Siegel & Dray, 2005). For example, we have found that German test participants expect a detailed written privacy policy as part of the main flow of screens at the beginning of a test. In contrast, U.S. participants would rather have a link to a separate privacy policy. Violating the expectations of either user group means participants will drop out from the study. URUT methodology must be adapted to different participant expectations in different countries, just like any other usability testing method.

The good news is that it is possible to find vendors who have done extensive URUT and other user testing in a host of countries (UXalliance, 2008) and can help you evaluate the important cost-benefit criteria for conducting global URUT.

Context of use is a defining element of usability (Bevan, 2008; National Institute of Standards and Technology, 2007). When considering the use of URUT, the researcher should determine whether or not participation can occur in the users' natural environment. URUT data can often generalize to real-world usage scenarios because participants are in their natural environment. Most participants can use their home PC and browser and can participate at any time during the day that is convenient to them. If the natural context of use is the participant's office, there could be some issues with firewalls or workplace policies. Make sure to check with your vendor about these issues if corporate workplaces are your target context.

Another practical issue when you consider URUT is that participants can be anywhere throughout the world, but they need Internet access with sufficient bandwidth to support your testing goals. For example, if you want to measure task completion time in a usability study via URUT, then you should be aware that the Web site or prototype you are testing can add some load time (and load time variability) to the pages you are testing. If you are performing a test in which timing is less important and/or stimulus complexity is low, such as a simple "where would you click to find X" test using static page prototypes, then potential network delays are usually not an issue. The URUT tool might add load time as well, for example, when the URUT server supporting your test is in a different country than the test

participants. However, most of the time the URUT tool will add no noticeable delays. Nevertheless, it is important that the load times are checked in advance. For one project, preliminary testing revealed page load times of 30 seconds when the prototype site was very heavy in flash and graphics and the URUT tool Web server was in a different country than the intended test participants. Testing ensures that the combination of URUT tool, participant bandwidth, prototype or site complexity, and testing goals is in harmony. Ask your vendor to provide data on network delays for each country you are testing in and, where you can, try to experience the delays yourself so you get a sense of what they are like qualitatively. A number describing network delays such as “5 seconds” may feel different than it sounds.

Large samples

Detecting statistically significant differences between groups or conditions is not usually possible with the small samples of a standard usability test (e.g., 10–20 participants). The large samples of participants used in URUT (e.g., more than 200 participants) provide you with enough statistical power to detect even small effects. For example, a 5-second difference in average task completion times (14.3 vs. 19.0 seconds) in a group of 200 participants will probably be statistically significant in URUT (depending on the standard deviations). However, with 12 participants in a standard usability test, it is unlikely that this size difference would be statistically significant. A full discussion of statistical power is beyond the scope of this chapter. Several good overviews can be found online, and Cohen (1988) is considered a good starting point. In short, statistical power is the probability that you will detect a difference in some measure between groups or conditions when that difference does, in fact, exist.

We have used 200 participants as an example here. If the differences that are meaningful to your project are larger than those in our example, then fewer participants could be used. If assessing statistical power for different measures is beyond the capability of the project team, make sure that internal or external consultants can help you determine the correct sample size to use based on the research goals.

The large samples of participants used in URUT give you not only statistical power but also a rich source of qualitative data because participants can provide typed responses to open-ended questions, just as they do in a survey. The qualitative responses can be grouped thematically to get a relative idea of how many participants express positive or negative opinions in certain areas, such as “pricing” or “search.”

It is always a good idea to have done moderated testing of some kind (e.g., usability testing and focus groups) to discover issues that are important to the research audience and inform the design of open-ended questions for URUT. Likewise, issues uncovered during online testing can be explored further in moderated sessions.

6.5.4 URUT in practice

All elements of a usability test script are embodied in the URUT tool. This includes instructions to participants; task wordings and order; randomization or counterbalancing schemes; rules for terminating a task; rules for accepting or rejecting participants based on an initial questionnaire; task feedback (e.g., “correct” or “incorrect” solution); and presession, post-task, and postsession questionnaires.

In this section, we discuss recruitment, pilot testing, different forms of testing, and the results and data. The following are the different forms of unmoderated usability testing:

- Task-based usability testing of a Web site or prototype, in which a test script is predefined before users are invited to participate
- True intent usability testing, which lets participants follow their own goals with a particular site or across many sites over an extended period of time
- Free search usability testing, in which users start from a blank page and we give them a specific goal to accomplish

Each of these forms of unmoderated usability testing is discussed in detail later.

All tools support task-based and free search usability testing, but many URUT tools also support true intent testing. A list of tools that support both moderated and unmoderated remote usability testing can be found at <http://remoteusability.com/?p=17#content>. Each vendor would be able to tell you how or if it supports task-based, free search, true intent, or other types of testing. Major vendors include LEOtrace, UserZoom, ClickTale, Keynote WebEffective, m-pathy, RelevantView, Chalkmark, VuLabs, SMT, ClickHeat, and Userfocus.

Recruitment

The URUT tool is typically agnostic to the method of recruiting participants. Anything you can do to point people at a URL is all that is required. Thus, you can invite people to participate via a visible layer on your existing site, just like many surveys do. If participants

click “yes” to participate, then they are directed to the tool, and the test begins. (The number of participants who opt-in varies widely.) Cookies can be used to prevent users from taking the survey multiple times.

The first step in recruitment is typically a screening questionnaire. If potential participants do not pass the screening, either because they do not match the desired profile or because the profile they match is “full,” then they exit the study. They also exit the study if they do not accept the privacy agreement or for technical reasons, such as having the wrong browser. Quotas for different types of participants can be set studywide (e.g., does not matter what country the participant is from) or per country, and some URUT tools help you manage and/or monitor these quotas.

Pilot test in each country

Some issues may loom larger for global testing via URUT than via traditional methods, such as the need to carefully pilot the test in each country to make sure that participants understand what is expected of them and that the test flows smoothly. Unlike face-to-face testing, there are no opportunities to make changes on the fly if participants seem confused about the response requirements. Therefore, pilot testing is even more important than when using traditional methods. How do you evaluate how well participants in different countries understand the flow and expectations of the testing? One answer is to hold a moderated session with some pilot participants in each country, either physically or via an online moderated session, and check for understanding of all elements of the test. You would ask questions such as the following:

- What do you think this task is asking you to do?
- Tell me about what you would type into this text box?
- What does each of these rating scale questions mean to you? Explain in your own words.
- What would a high rating mean? What would a low rating mean?

If you or your vendor(s) has done this type of testing before, pilot testing at this level of detail may not be necessary because the requirements for different countries may be well understood, especially if the questions and other stimuli have been used before. Anytime new questionnaire items and task wordings are used, it is best to validate those.

An alternative style of pilot test is to use the exact same URUT method over an internal sample of users prior to launching invitations to the

target users. A set of 10 invitations can be generated for those internal users (client's team members in different locations, project managers, researchers, and potentially even friends) to try to test the test protocol remotely before launch. In global research, it is very useful to test wording, language, and other cultural issues, again with the advantages of participating remotely. Afterward, editing the protocol is quite easy, so the entire pilot study can be done within a fairly brief period of time, even if one needs to test in various countries. There are risks to including only internal users in a pilot test. For example, there may be terminology that is only understood within a company and not by external end users.

Task-based usability testing

Before tasks are presented, participants read instructions about how the session will flow and how to give feedback. Different profiles, including different countries, can be given different tasks. If desired, tasks can also be presented conditionally, based on some performance criteria such as success on a previous task. Questionnaires can also be presented conditionally. Other types of rules can be used to control the flow of the study. In one test, to keep the test from taking too much time and to keep participants from getting frustrated, tasks were stopped if participants made more than five navigation clicks. A task can also be stopped if the participant “does the right thing,” which could be clicking the desired button, reaching the desired page, or completing a field with the proper format. The tool can then automatically tally a 1 or 0 in the success column so that you can calculate some measure of “effectiveness” for quantitative usability. Do not use more rules for presenting questionnaires and controlling the flow of the test than you need, however. Later, we discuss how this adds to the cost of the test via increased programming and testing time.

Figure 6.3 shows an URUT tool presenting a task in a frame at the bottom of a browser. The participant can indicate the end of the task by pressing one of the buttons (Success or Abandon). Buttons can be combined with rules to terminate a task. Branding is often an optional element of the task bar.

True intent usability testing

Sometimes it is better to observe what users actually do on a site and ask them about their goals rather than provide them with predefined tasks as you would for a standard usability test. This is called true



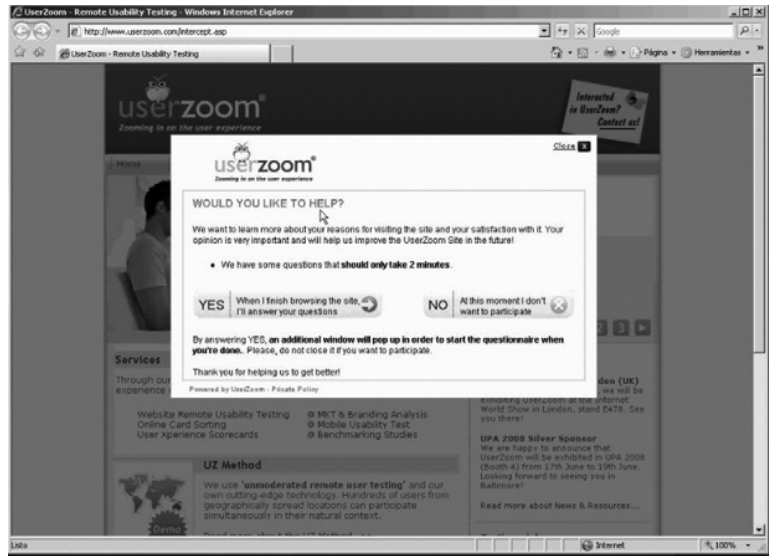
■ **FIGURE 6.3** A live Web site being tested via an URUT tool.

intent testing. The same URUT tool that supports usability testing can typically be used for true intent testing. True intent testing helps you answer questions such as the following (see also iPerceptions, 2008):

- What are the goals of a user who comes to a certain Web site?
- For each user goal, what are the success and satisfaction rates?
- What features and paths are being used for each goal?
- Why do participants do what they do?

Once the participants are recruited, they complete a demographic survey, are asked what they came to the site to do (i.e., what is their goal), accept a privacy agreement, and read instructions on how to give feedback. This is similar to the procedure for task-based testing.

However, rather than being presented predefined tasks, they are asked to do whatever it is they came to the site to do that day (Fig. 6.4). Rather than task descriptions being presented at the bottom of the browser (see Fig. 6.3), a “feedback” button is provided so participants can provide spontaneous feedback as they use the site and a “finished” button



■ FIGURE 6.4 A true intent study starts by asking a visitor to participate.

to indicate when they are done giving feedback. Their session can end when they press the button or when they exit the site, or some other trigger can be used, such as amount of time on the site, page visited, or number of clicks. At the end of the session, they fill out a questionnaire asking how successful they felt based on their stated goal and their likes and dislikes about the site. Participants also provide ratings on several questions.

True intent testing does not require you to translate task descriptions for each country. You still need to translate the initial and final questionnaires and task instructions. However, if you or your vendor has already done this once, then each time you collect data, you only need to translate verbatim responses. Thus, this type of global testing can be quite economical when used repeatedly – for example, to do quarterly evaluation of Web site “status” in each country. True intent testing poses different expectations about giving feedback, especially verbatim spontaneous feedback while the user surfs, so pilot testing should be done at least once in each country to ensure that participants understand what is expected.

Note that effectiveness and efficiency may not be what you want to measure in a true intent study, but subjective satisfaction always applies. Subjective satisfaction and perceived success compared to

stated goal and path taken will be important quantitative measures. The spontaneous feedback during surfing and verbatim likes and dislikes will be the important qualitative data.

Free search usability testing

A free (or open) search usability test consists of asking users to complete a specific goal starting from a blank browser page. In other words, instead of having users start from a specific Web page and evaluate its usability, users start from scratch. In this case, the research objective is to analyze online behavior (and not so much a Web site's usability).

Imagine that you own a car accessories Web site, aimed mostly at sports car owners. For example, one of your specialties is high-quality car cleaning products for delicate paints. Wouldn't it be interesting to know how sports car owners search for these types of products online? Would you be interested in knowing what terms users entered in order to find them? How about the sites visited? Did they choose a particular brand or Web site, or did they use a search engine? Most important, why do they search the way they search? Are they looking for specific information? If so, what kind of information?

Through URUT, you would be able to perform a free search study and obtain the answers to these questions. In the free Web search study, you could ask 300 sports car owners to start from a blank browser page and look for car cleaning products for delicate paints. The instructions of the participant might be as follows:

You own a red Corvette (or similar) and you'd like to keep it looking great every other weekend by hand-washing it yourself. You know it takes a special kind of cleaning product to help you wash the car without scratching it, so it looks and shines like new. Based on this scenario, try to find a product that would satisfy your current needs. You will start from a blank page and you may freely go wherever you want.

The test results would help you to understand customers' behavior in a way that will help you better position your site on the Web. You will also obtain marketing and branding data that will help you in your search engine marketing strategy. You will know if competitors are being recognized as well. Finally, you will better understand your customers' mental models and what terms they use so you can make sure your site will fit in with those.

Results and data

One of the main reasons why URUTs are being implemented quite successfully in a growing number of firms conducting usability testing and user experience research is the quality of the data collected. It is, most of all, actionable and relevant data. Here is a list of results gathered:

- Effectiveness and efficiency ratios – rate of users who completed a task successfully, plus the time and number of clicks required to complete the task (Fig. 6.5).
- Satisfaction – how satisfactory was the experience.
- Answers to the specific questions – there are initial, introductory questionnaires, task-specific questionnaires, and final, overall satisfaction questionnaires. Through these, various types of questions can be asked, such as multiple-choice questions, Likert scale ratings, and open-ended questions (for user verbatim) (Fig. 6.6).
- Behavioral results – “click-streams” show the navigation path users followed to complete a specific task. “Click maps” show the aggregate sum of where all users clicked (Figs. 6.7 and 6.8).

In addition to these results, it is very important to note that URUT tools usually offer many scripting and filtering or segmentation capabilities that allow researchers to fully customize their studies and perform advanced data analysis. For example, a researcher may want to analyze what actually happened to users over a certain age group and from a specific country who failed task number 2, what path they



■ FIGURE 6.5 Effectiveness and efficiency ratios.

Final Questionnaire									
1. Please, answer the following questions concerning the web site, taking into account that 1= Completely disagree and 7= Completely agree.									
	Answer distribution							N	Mean*
	1	2	3	4	5	6	7		
The Web site clearly explains what it is about.	1% (2)	2% (5)	6% (17)	10% (31)	24% (72)	30% (90)	28% (83)	300	5,6
The Web site is attractive.	1% (2)	1% (2)	2% (7)	6% (19)	20% (59)	38% (113)	33% (98)	300	5,9
I can clearly see the products and services the Web site is offering.	1% (4)	1% (4)	4% (11)	7% (21)	17% (50)	34% (103)	36% (107)	300	5,8
The Web site seems easy to navigate.	1% (3)	1% (3)	2% (5)	3% (9)	19% (58)	33% (100)	41% (122)	300	6
Important information is featured more than secondary information.	0% (1)	1% (3)	4% (11)	8% (23)	27% (80)	29% (87)	32% (95)	300	5,7
The different sections of the Web site are easy to view.	0% (1)	2% (5)	2% (5)	6% (19)	19% (58)	34% (103)	36% (109)	300	5,9
The Web site gives me the feeling that it is reliable and trustworthy.	1% (3)	1% (4)	3% (9)	9% (26)	19% (58)	36% (108)	31% (92)	300	5,7

*(1- Completely agree / 7- Completely disagree)

FIGURE 6.6 Likert scale ratings.



FIGURE 6.7 Click streams.



FIGURE 6.8 Click maps.

took, and only the ones who spent more than 2 minutes on the task or even the ones who did not see a specific page. All this can be done by using the filtering capabilities of URUTs.

6.5.5 Technical considerations

Technology constraints on data collected

In general, anything that can be accessed via a URL can be studied with the URUT tools. However, technology constraints may limit the information you get, depending on the tool you choose. In some cases, navigation cannot be tracked if URLs do not change – for example, if a Web application makes database calls to refresh parts of a screen (e.g., dynamic Web technology such as AJAX). To track interaction with video elements, special programming may be required, which can affect study cost. For Web applications that do not have an externally visible IP address (i.e., they are behind a firewall), the URUT tool might not work at all. It is best to check these technology issues when selecting the correct tool for your study.

Browsers and operating systems

The URUT tool can check the compatibility of operating system and browser and selectively screen participants based on their configuration. Most URUT tools support the Windows operating system and the Internet Explorer browser. Some tools also support Mozilla browsers. Compatibility with other browsers or operating systems is something you can test but should not be assumed.

Software download and installation

For most URUT studies, there is no software download or installation required of participants. Participant interactions are routed through a proxy server that records them and forwards them to the site's server. The proxy server approach works if the URLs to be evaluated are known in advance.

However, if you want to study behavior across several Web sites that are not known in advance (e.g., in true intent studies), participants will need a browser plug-in. The effort for participants to download and install a browser plug-in is fairly small and follows the same process as for any browser plug-in. The requirement to download a plug-in can have a negative effect on the participation rate for the study. The instructions that you provide to participants for downloading and installing the plug-in will have an effect on their willingness to participate.

The browser plug-in approach is also used for longitudinal studies (e.g., studies in which the same participants are observed repeatedly over a period of time). For example, we studied 800 participants over a period of 4 weeks. They were recruited based on the fact that they shop online. They installed the browser plug-in for either IE or Mozilla, and we recorded the Web sites they visited. In addition, surveys were presented when participants exited any one of 120 sites on our “interesting sites” list. The plug-in also controlled how many total surveys participants received for the entire study, per week and per day. As you might imagine, 4 weeks of navigation and survey data from 800 people is a very large and rich data set. There were more than 3 million data records. This amount of data cannot be opened in Microsoft Excel. We wrote custom scripts (in Perl, although other languages such as Ruby would do just as well) to summarize data and correlate behavior and opinions. A usability test that occurs in one day will not generate that much data and will most likely be analyzable directly in Microsoft Excel. However, scrolling and navigation data lend themselves to visualization not provided in Excel.

A longitudinal true intent study can be used to capture best practice in a product category and answer questions such as the following:

- What are the characteristics of sites that attract positive ratings and comments?
- What sites are visited together, in what order, to support which goals?
- How do the previous issues vary by product category (e.g., electronics and books)?

6.5.6 Cost and benefit in global URUT

The URUT approach to user research allows you to test large numbers of participants, which provides statistically powerful data from a global audience participating in their natural context and often using their own goals. Quantitative benchmarking between sites or between countries is therefore possible. Qualitative issues can be both discovered and quantified. Traditional usability testing allows you to interact with participants and probe on qualitative issues in a way that cannot be done with URUT. URUT provides quantitative measurement that cannot be matched by traditional testing. A sequence of traditional testing followed by URUT would be ideal. The traditional testing would allow a deeper exploration of qualitative issues in each country and would inform the URUT, which would bring quantitative

power to the testing. However, in practice, one method will likely be chosen over the other based on the relative costs and benefits in the context of the study goals.

There are three unique costs associated with URUT that do not apply to traditional testing. First is the cost associated with the requirement to perform quality testing and regression testing on the URUT software. The testing effort increases exponentially with the number of rules that are programmed, such as counterbalancing, task termination rules, and situational questionnaire presentation. Based on our experience, you should keep the test as simple as possible. Adding many “nice-to-have” rules can increase the software testing effort and cost beyond the value provided by the additional complexity of the test. The other two costs that are unique to URUT are associated with the additional upfront planning required of an automated test and the need to program the automation. Thus, three additional costs for URUT testing are related to planning, programming, and testing the automation.

These additional costs are typically more than offset by cost savings. First, travel costs are eliminated with URUT. Additional staffing costs per country are also eliminated. The programming of rules and questionnaires (not the content) applies to all countries, so the incremental cost per country is very small. Cost savings for URUT really kick in when you repeat testing with the same or substantially similar setup. In the case of true intent testing at a regular interval (e.g., quarterly), there are no or low incremental costs associated with planning, programming, and testing. A before/after usability test also has savings from reuse of elements.

Translation costs, for both test materials and verbatim data, are similar for traditional testing and URUT.

6.5.7 Case study: Global testing of Monster.com

Monster.com is a well-known Web site on which home users can search for jobs and professional users can post job openings (Fig. 6.9). It is a true global brand and .com, with local Web sites in 50 countries (Monster.com, 2009). As such, it required local user experience and usability research as part of its global strategy.



■ FIGURE 6.9 Monster.com.

Situation and challenge

In 2007, Monster.com's interface was about to be redesigned and the firm's user experience team was searching for a way to perform global testing in various markets. The challenge was a relatively low budget and short time frame.

To get an accurate and objective assessment of the user experience internationally, Monster conducted user testing with more than 1000 users using a URUT solution in its four main European markets (United Kingdom + Ireland, France, The Netherlands, and Germany).

Solution

Monster.com successfully implemented the test using UserZoom's URUT solution. The study included 250 users per market (1000 total), five tasks to be completed, plus the questionnaires. Recruited users were segmented by geographic location, age, sex, home users who searched for a job, and professional users who searched for candidates.

Results

The report generated included the following:

- Effectiveness, efficiency, and satisfaction ratios
- Click streams and click maps
- User verbatim and suggestions to improve
- Overall user experience rating and opinion of the site about various issues, such as graphic design, quality of the information, and legibility – all done post-task
- Advanced analysis of the data gathered, including data segmentation by specific variables, cross-tabulates, and comparison between markets

The tests revealed an abundance of navigation and labeling problems that needed to be fixed, especially in the two most important parts of the site: job search and the site's CV builder.

6.6 WEB ANALYTICS

Written by Adam Cox and Tjeerd de Boer

6.6.1 Introduction

Most user research professionals have heard about Web analytics, even if they are not regular users. In this section, we discuss the use of Web analytics as a method for global user research. A logical starting point, therefore, is to determine what Web analytics is. Web

analytics is officially defined as “the measurement, collection, analysis, and reporting of Internet data for the purposes of understanding and optimizing Web usage” (Web Analytics Association, 2008). However, what does this actually mean?

Measurement – Basically, all visitor activities on a Web site can be measured (e.g., mouse clicks, data input, referrals, page views, and number of visits). In addition, it is possible to determine how the Web site was accessed (e.g., search engine, search terms, referrals, and bookmark), the technical characteristics of the device that was used to access the Web site (e.g., browser, operating system, and screen resolution), and the location of the user (e.g., country, city, and region within a country).

Collection – This refers to the act of gathering and parsing Internet data into a format readable by humans. This is performed by Web analytics software. There are many different Web analytics software solutions (e.g., WebTrends, Google Analytics, and Omniture), but there are only a few different collection methods (e.g., server-log based or tag based are the most common).

Analysis – Internet data need to be interpreted by a human. The numbers alone do not give any meaningful insights. Analysis is normally carried out by a Web analyst, but the basics might be able to be done by a marketing or user research professional.

Reporting – A Web analytics report can range from a basic monthly statistics report that reflects the performance of a Web site to more complex and detailed reports that are used to answer specific questions. In any case, the assumption is that the report will be used to communicate findings and possibly to suggest improvements.

Understanding and optimizing – Both user research professionals and Web analytics professionals want to understand how people are using interactive technology so that they can improve the products and services.

Web usage – Unlike the broader scope of usability testing, Web analytics is not about all interactive products such as parking meters or DVD players. It is about Web applications or sites that can be accessed with different devices (desktop computer, mobile device, etc.).

Web analytics is sometimes perceived as a technical issue. Indeed, it takes some technical knowledge to implement Web analytics software and start the measurement and data collection. However,

this technical work can often be done by people without a high degree of technical knowledge (internal technical staff, software vendors, consultants, etc.). Once the software is installed and functional, Web analytics is primarily concerned with nontechnical issues. For example, one needs to determine the appropriate research questions (e.g., who is visiting the Web site) and select the Web data that can answer these questions (country, operating system, return visitor, etc.). In addition, data need to be analyzed and reported. The analysis, therefore, should be done by people with an understanding of the target audience, the product, and the business goals. Considering these characteristics, it is easy to see that user research professionals fit this profile. Of course, some basic knowledge is necessary to succeed, although the recent availability of easy-to-use and inexpensive (sometimes even free) Web analytics software has removed many of the boundaries to participation. In other words, any research professional is able to do basic Web analytics activities. The more complex activities (A/B testing, multivariate testing, behavioral targeting, real-time customization, integration with CRM systems, etc.) should probably still be done by full-time Web analytics professionals. Therefore, this section addresses only basic Web analytics activities that are relevant to user research professionals. Information about more complex Web analytics, solution providers, and useful resources can be found on the Web site of the Web Analytics Association (<http://www.webanalyticsassociation.org>).

Why, then, should a user research professional be concerned with Web analytics? Because Web analytics takes a truly global approach to answering research questions; it transcends boundaries and is therefore truly global. Web analytics can be used to monitor user behavior throughout the world, even if the Web site of interest is local.

In the remainder of this section, we discuss the combination of Web analytics with global usability testing and illustrate how to execute this practically.

6.6.2 Why combine Web analytics and global usability testing?

Introduction

Combining user research methods is a valuable practice. Whether it is the combination of eye tracking and usability testing or Web analytics with an expert review, the benefits are obvious. The combination of

Web analytics with global usability testing is a particularly powerful approach because although the methods are very different, they complement each other well. Combining Web analytics and usability testing allows the user research professional to utilize the benefits of each method, creating a much more valuable report. This is even more important when doing global user research because the variety of visitors and behavior is often overwhelming.

In the following paragraphs, we highlight the main benefits of combining Web analytics and global user research. We do this by comparing Web analytics with usability testing and showing how they complement each other. Other user research methods can be combined with Web analytics as well (e.g., online survey and expert review). However, because the methods of Web analytics and usability testing are so different, their combination can be very successful.

Scope

Usability testing has few limitations in terms of what can be evaluated. For example, paper (prototypes), interactive display booths, or MP3 players can all be evaluated during a usability test. In contrast, there are more constraints associated with Web analytics. The interactive product must be a Web site, it must be “live,” and it must generate a significant amount of traffic. In addition, to be able to monitor the Web site, Web analytics software must also be implemented and set up properly. It is quite common, for example, for Web analytics software to be installed but not properly configured or used to optimize a Web site. Implementation of Web analytics software can be complex and should never be underestimated.

Measurement

Web analytics is, to a large degree, a measured method of evaluation. This means that it is reliable and objective. The reliability of Web analytics data is also determined by the quantity of site visitors. Some Web sites attract thousands of visitors, perhaps tens of thousands, every day. With large numbers, it becomes very unlikely that measurement errors can influence the final results. Web analytics also uses relatively standard definitions for different measurements. Unique visitors, return visitors, and bounce rates, for example, are established terminology and can be used in a standard, quantitative way, which makes it easy to compare. In contrast, usability testing is much more subjective – observations are done by humans – and therefore considerably more labor-intensive.

Compared to global usability testing, Web analytics offers another benefit in that it can be more efficient. With Web analytics, you can evaluate Web sites with users throughout the world efficiently. This is very time-consuming and expensive when carried out with traditional usability tests. Compared to other user research methods, Web analytics can be relatively inexpensive and automatic (given the assumption that it is installed and set up properly). However, this does not mean that Web analytics is always inexpensive. When more complex Web analytics activities are carried out (e.g., path analysis, A/B testing, and behavioral targeting), the costs will increase significantly because full-time Web analytics professionals are often necessary.

Validity

In general, Web analytics has higher validity compared to usability testing. This is because Web analytics monitors the following:

- Actual use (what users do versus what they say)
- All tasks (which can be more realistic and reflects users' motivation)
- All sessions (longitudinal aspect and continuous monitoring)
- All users and all behavior

However, Web analytics is limited to monitoring Web behavior only. Usability testing, on the other hand, can evaluate a user's emotion and behavior when using products and services other than Web sites.

Insights and results

Usability testing can give excellent insights into the *why*. These insights come from direct observation and interaction between the moderator and the test participants (e.g., emotions, feedback, and motivations for behavior). A usability test can facilitate a greater understanding of the context of the target audience, especially when a test is conducted on location. In addition, sometimes a single test participant can give you more results than a complete Web analytics report.

On the other hand, Web analytics can yield insights that are not normally gained by usability testing. For example, Web analytics can give insight on where users land on the Web site when they have used a search engine to find your site. In addition, Web analytics makes benchmarking and comparisons much easier because the Web site can be continuously monitored and evaluated between iterations and over time.

6.6.3 Combining usability testing and Web analytics

The combined approach of both methods is usually better than a global usability test alone. This section discusses the practical side of combining the global usability test and Web analytics. Ideally, Web analytics activities should be conducted before the usability test starts. A Web analytics report, for example, can be used as an indication of problem areas that need to be included in the usability test. In general, we distinguish five phases of the global usability test:

1. Planning and preparation
2. Conducting the sessions
3. Analysis
4. Reporting
5. Using the results

Web analytics can be relevant during all phases *except* when conducting the test sessions (phase 2). The following discussion focuses on using information from a basic Web analytics report in the different phases of a global usability test.

Planning and preparation

Test goals

It is likely that Web analytics activities have uncovered a number of problem areas. Some of these problems will be well understood and can now be left out of the usability test. However, other problems may be less clear, and usability testing can be used to get a better understanding of why the problems exist. In addition, the usability test should focus on finding problems that are difficult to identify by using Web analytics. The Web analytics report might show that 40% of the visitors leave the Web site within 1 minute. If this is considered a problem, then the usability test could be used to find out why people are leaving the Web site.

Participants

Web analytics allows the user research professional to discern profiles for the visitors of the Web site in question. This can be used to determine the appropriate recruitment profile for the participants. The following are examples of data that can be used to inform the creation of participant profiles:

- Investigating users' geographic location – whether they are local, national, or international visitors. This is especially important when considering the global aspect of usability

testing because it could provide conclusive evidence about which countries are the most important.

- The time of day and days of the week can help to determine the context of visitors – that is, home user versus business user.
- Type of Internet connection, screen resolution, browser, and operating systems are all standard metrics that can bring more context to understanding user groups, especially when compared to the general Internet population.

Web analytics can also identify different segments of users by examining trends and patterns of behavior. For example, visitors from one referral might be more likely to purchase a product on your Web site than visitors from another referral, or visitors from a certain geographic location may be more likely to be interested in certain content.

Finally, Web analytics can be used to determine the gap between the target audience of the Web site and the actual visitors. If there is a large gap, this can lead to interesting research questions for the usability test.

Selecting test tasks

Web analytics can be used to define tasks that are given to the participants during the test sessions. For example, it can give an indication why users are visiting the Web site and (sometimes) why they are leaving. This is done by the following:

- Looking at search terms and key words that visitors used to reach your Web site in order to gain insight into what they were looking for and their intentions.
- Investigating on-site search functionality to provide an indication of what visitors might not be able to find using other methods of navigation.
- Investigating referrals to understand where some of your users are coming from and if they match your target audience.
- Looking at the *bounce rate*, which is one of the most important indicators of a Web site's success. It effectively represents the percentage of visitors who land on a Web page and leave without viewing any other pages. Depending on the nature of the Web site, an acceptable bounce rate is often between 30% and 40%. Anything higher, especially for a home page, raises serious questions. If the bounce rate is very high, the Web site can be tested during a usability test to determine why visitors are exiting the site so quickly.

- Understanding which pages generate the most visits is important for determining where they are coming from (e.g., entry and exit pages). For example, you might discover that most of the visitors do not land on your home page, but through a key word search on Google they land on a product page. This would help prepare for usability testing because you have an idea of where to start the test. Similarly, exit pages provide useful insight. For example, an exit page could be as common as a “thank-you” page at the end of a checkout process or something less predictable, such as a content page. Either of these two types of pages can be isolated for further optimization or, indeed, provide a focus for usability testing.

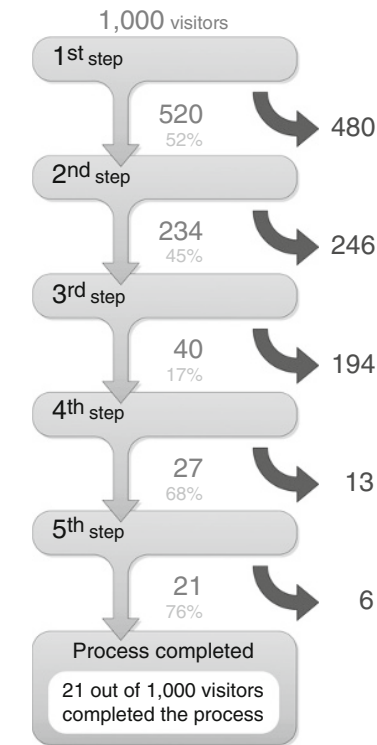
Web analytics can also show what visitors are doing on a Web site, which can be compared to what would be considered normal or desired behavior. This is done by path and funnel analysis. Path analysis is a technique to look at some of the most common user journeys through a Web site. Insights can be gained as to what users are actually doing on the Web site compared to what they should be doing. A similar technique, called funnel analysis (Fig. 6.10), is used to investigate a predetermined user flow (often a checkout process with many steps) to find the main problem areas and identify the conversion rate (2.1% in Fig. 6.10). Both techniques can be difficult to investigate but can be very useful when preparing for a usability test because they will provide focus.

The insight into why visitors are coming to the Web site, why and where they are leaving, and what happens between those two events can make it easier to select the proper tasks for the usability test.

Research objectives and environment

Web analytics can show what parts of the Web site were used most often, how visitors got there, or what functionality gave the most problems. The research objectives of the usability test can therefore be limited to these aspects and scenarios.

In terms of the test environment, Web analytics can give substantial information on the context of the visitor (e.g., access from home or from work, single session versus multiple sessions, operating system, browser, Internet connection, and screen resolution). In most cases, the test environment should reflect this context and should therefore be simulated in the test lab.



■ **FIGURE 6.10** Example of funnel analysis.

green visitors advancing to next step
 red visitors “abandoning” process
 grey percentage of visitors advancing to next step

Analysis

When combining Web analytics and usability testing, it is possible to compare the findings for verification. If the findings offer similar conclusions, the Web data can make the analysis of the usability test easier. For example, the Web data can be used to determine the severity of a problem (e.g., do visitors leave the Web site after experiencing a specific problem?).

If, on the other hand, the findings do not point in the same direction, this can be interesting as well. Usability testing could conclude that participants claim to behave in a certain way, whereas the Web data could prove otherwise. An example of this is when participants declare their interest in a site and that they will use it frequently, but the Web data show that the percentage and absolute number of return visitors are actually low.

The analysis of the usability test results frequently leads to new questions. Referring back to Web analytics can often answer these questions. For example, suppose a usability test was carried out with 10 participants and only 1 was unable to complete a checkout procedure. Is it therefore appropriate to conclude that this is a minor problem? In this case, examining the Web data again will provide you an indication of the frequency of this problem. Of course, the same principle applies to global usability tests: If the analysis shows an outlier (whether this is a participant or a country), you can use Web analytics to get more information on the severity of this behavior.

Reporting

Web analytics and usability testing results should usually be combined into one report because this is more insightful and efficient than two separate reports. This report shows the findings of both methods and how they support each other. This combined report yields more certainty on the issues identified because it is based on two very different and complementary sources. For example, the results of a usability test with “only 12 participants” can be backed up with the Web data of thousands of visitors.

It is not just the increased certainty that makes the combined report more powerful; the numbers provided by the Web analytics seem to be especially attractive to members of higher management, as they are used to working with numbers. If you want to convince these decision makers with your results, you should speak their language, which means you should include numbers.

Using the results

Once possible solutions have been implemented, how do you know that the changes to the Web site will indeed be a solution? Although you can conduct a second usability test, this can be quite time-consuming and expensive. It might be easier to examine the changes in the Web data following the supposed improvements to the Web site. With appropriate benchmarks, Web analytics can give an indication of whether these changes made on the basis of usability testing improved the Web site’s performance.

In addition to measuring the impact of changes to a Web site, Web analytics should also be used to continuously monitor the performance. The continuous and real-time feedback that is provided can change an organization’s mind-set to one of continuous optimization. Also, if there are questions that Web analytics cannot answer, you can always

conduct another usability test. So it seems that the fields of Web analytics and usability testing have started to join efforts.

6.6.4 Trends

Web analytics, as a field, continues to evolve. Since its technical beginning of server log files analysis, all aspects of measurement, collection, reporting, and analysis of data have improved. The software has become increasingly powerful and easier to use, which makes it more interesting to the user research community. Interestingly, Web analytics professionals are also showing an increased interest in user research methods, especially usability testing.

Another trend is the appearance of numerous tools that can be placed between usability testing and Web analytics. See, for example, the section on unmoderated remote usability testing. ClickTale is software that records all mouse movements, clicks, and keystrokes of visitor's sessions, which can then be played back as video clips or summarized as heat maps. Unlike other well-known screen action capture tools that require installation on a computer, such as Camtasia or Lotus Screen Cam, ClickTale can be implemented on the Web site by adding tags to the source code. This goes a long way in terms of not just understanding what users are doing but also why they are doing it. Of course, watching mouse movements, clicks, and keystrokes does not tell you exactly why the user is exhibiting a certain behavior, but sometimes this is very obvious (e.g., the user fills out a form). Even when this is not obvious, you can still make assumptions and check these using other research methods (e.g., usability testing). No doubt, more and better user research tools will be developed in the future, but the existing tools are very useful already.

A final trend is that the acceptance and adoption of Web analytics within organizations is increasing. There are even organizations that could be called "data driven," such as Amazon, Google, and Netflix. These companies actually compete on analytics and are market leaders for a reason. They rely on analytics to make decisions. Their analytics are based on more than just Web data; they are closely integrated into their business and sales systems. In an increasingly competitive business environment, accurate and predictive analytics are required. However, these companies also need the insights that can only be provided by user research. Therefore, the future looks bright for both user research and Web analytics, and especially so for the combination of the two.

6.6.5 Conclusion

The combination of Web analytics and usability testing is a very valuable approach to evaluation. It is likely that you have some form of Web analytics available on your Web site, so it makes absolute sense to take advantage of it by combining Web analytics with another form of evaluation. If you are doing global usability testing, then these combined methods are especially beneficial. It can be used in the preparation of usability testing, validation of results, and in support of the process of continuous optimization. In other words, if you are not doing Web analytics yet, you should be starting now. If you are doing Web analytics already, then you should be combining it with global usability tests and other user research methods.

6.7 ONLINE SURVEYS

Written by Inga Luedemann and Torsten Müller

6.7.1 Introduction

During global user research projects, it can be very complex and time-consuming to coordinate face-to-face interviews in all targeted countries. You need local teams that manage recruiting, conduct interviews, and perhaps analyze the data and write the reports. Using surveys can be helpful in these situations, and they are often used for this reason. Surveys can be administered by mail or phone, but for global research this is usually carried out via the Internet (i.e., online surveys). These online surveys are the topic of this section.

A large variety of tools are available for creating, administering, and analyzing online surveys. Almost all tools support standard question types (e.g., open-ended and multiple-choice) and standard interaction techniques (e.g., radio buttons, check boxes, and drop-down lists). The main differences can be found in the more advanced features (e.g., randomization of questions or responses, skip logic, and data export). A list of popular online surveys can be found at <http://www.measuringux.com/linesurveys>.

6.7.2 Use of online surveys

Online surveys can be useful in almost all development phases of a new online product or service. Early in the development process, online surveys can be used to get feedback about desired features and functions or what is confusing about an existing application. During iterative design and prototyping, online surveys can give you

feedback about design alternatives. In these cases, a link to an online survey is sent to the participants. These participants are selected in a similar way as for usability tests (i.e., recruitment based on a profile).

After the launch of the online Web site, you can ask the actual visitors to give feedback using the online survey. The main advantage of this recruitment is that you can be sure that these participants are actual visitors of your Web site, with real needs and goals (“Why are you coming to the Web site?” “Did you find what you were looking for?”). The main disadvantage is that the participants might not be representative for the entire group of visitors. This is, for example, because people with a strong positive or negative opinion can be overrepresented in the group of participants. In addition, users who visit the site for the first time are often underrepresented. Finally, the group of visitors might not be the same as the target audience, making it impossible to ask why people are *not* coming to the site.

There are several ways to initiate the online survey to visitors of a site. Usually, the survey is presented when visitors arrive on the site or are leaving it (“Please give us a few minutes to answer these questions.”). A combination of these two ways is also possible: Visitors are invited when they see the first page of your site and the survey is presented when they exit the site. Finally, an online survey can also be initiated during the sessions. This approach is relatively new and gives you very interesting possibilities because the survey can now be triggered by your Web analytics tool. Specific groups of visitors (e.g., returning visitors) or visitors displaying specific behavior (e.g., customers who did not complete a purchase) can be asked to participate. Most Web analytics tools do not support this integration with online surveys yet, but this will probably change in the near future.

6.7.3 Global online surveys

Online surveys have a number of characteristics that make them well suited for global research:

- Global online surveys are usually conducted by a single agency and therefore can be run very efficiently in terms of time – and hence costs – because there is less effort needed for coordination. The role of local agencies is usually limited to the localization and translation of the survey questions.
- You reach a large group of people in a very short time (which is again cost-saving), and you also have access to different kinds of people very easily.

- Participants are not influenced by an interviewer or a lab environment because they can answer all questions and tasks at home or at work or wherever they feel comfortable.
- If the survey is adequately designed and implemented, the data entry should be fairly easy. Programmed filters and restrictions should prevent participants from making mistakes, which has a positive effect on the quality of the data.
- Finally, the data are available directly after the field phase and can be analyzed immediately.

These are the main reasons why online surveys are useful for global user research. It is important to note that surveys are typically used as a supplementary method. In other words, they are used alongside other user research methods and not as the only research method – for example, as a preliminary survey, as a validation of specific qualitative research questions, or simply as an efficient data entry method to support face-to-face interviews.

Global online surveys show some clear benefits, but they also have a few disadvantages:

- An online survey is a very restricted way of getting results because it must be simple and relatively brief.
- In general, with online surveys it is less obvious that the person completing the survey is really part of your targeted participants. Therefore, you need to make sure to address the correct audience.
- It is difficult to be sure what the true motivation is for participants to participate and how motivated they are to answer honestly. Unlike face-to-face interviews, it is impossible to follow through on a given answer.
- During completion of the survey, participants may abandon the survey for all sorts of reasons. It is difficult to determine the reasons behind abandonment and impossible to motivate participants to continue.
- Global online surveys need to be translated and sometimes adjusted for specific local use. This can be an expensive and time-consuming process if not managed adequately.

Despite the challenges, conducting a global online survey need not be overly complicated. From a technical standpoint, you simply need a multilingual survey hosted on a Web server that is accessible from all relevant countries. However, there are a few guidelines and issues to bear in mind to ensure the success of your survey. These are addressed in the following sections.

6.7.4 Short and simple surveys

One of the main challenges in designing online surveys is to get a high response rate and genuine answers. Therefore, you need to make sure that the survey can be completed quickly and easily by the participant. It also helps when the participant believes that he is taking part in a survey that has meaning to him and for which his responses may actually make a difference.

The design of an online survey should be managed just like any other user experience project. It should be made easy for the participant to complete her task. Therefore, the survey should be short and easy to understand.

Rather than asking too many questions of your participants, consider asking different questions of different participants. In this case, you would have to randomly assign one of the survey versions to each participant.

6.7.5 Scales

Another challenge in designing global online surveys is the use of scales. Scales that work well in some countries might be interpreted differently in other countries. This is especially true for scales that use school grades. School grades are a good option to use in single-country studies because participants will be familiar with them from their own schooling and therefore should not have difficulty interpreting the letters of the scale. However, most countries have unique grading systems, so an English participant will use the German grading scale differently than a German participant. Thus, avoid scales consisting of school grades in global projects. Moreover, try to support the numbers or letters you use for scales with visual elements, reducing bias on results arising from translation issues.

6.7.6 Review

The quality of the survey is essential for getting the correct results. Therefore, surveys should always be reviewed prior to administration. This is even more important for global surveys. Because the survey has to be translated into different languages, it is crucial to get agreement and sign-off on the survey before it will be translated. The costs of translation are generally high, and it becomes difficult and costly to make changes once the survey has been sent to the translation agency. The best way to approach this is to have some initial reviews before programming the questionnaire in one language. Once the programmed questionnaire has been approved by all stakeholders, it can be translated into different languages.

6.7.7 Professional survey software

When developing an online survey, especially a large global survey, it is recommended to use professional survey software. In general, this allows you to program a general questionnaire structure (question type, number of answering alternatives, scales, filters, etc.) and use separate text files as source documents for each language version of the survey. Thus, changes regarding the structure of the questionnaire have to be applied just once, which significantly reduces the effort and the probability of mistakes. Moreover, data for all countries will be delivered within one data set, saving you hours of matching data from several countries or writing separate tabulation syntax for each country.

6.7.8 Manage contractors

When conducting a global online survey, it is likely that you will be working with international contractors. If possible, use just one translation agency for all the languages into which you need the survey to be translated. The same holds true if you are seeking a provider for survey panels. This will reduce your communication efforts toward translators because you have just one contact for all translations, which in turn will save you much time and money because you do not have to provide comments and explanations for each language. The agency should use professional translators and have experience within the area of market research. Domain-specific experience will allow the translators to be more accurate and efficient in their translation of the survey because they will already be familiar with the appropriate terminology.

6.7.9 Conclusions

Online surveys for global user research are best used as a complementary research method to other methods. These can be a relatively cheap and efficient way to gain additional insight into your global user base. Online surveys can be used during almost all phases of a development process, including after a Web site has been launched. When used for an existing site, online surveys can be presented when visitors enter the site, when they leave, or at some point during the sessions. Web analytics tools can be used to ask specific questions to specific groups of visitors rather than asking the same general questions to all visitors.

A few guidelines need to be kept in mind when conducting global online surveys: The survey should be kept short and easy, the

questionnaire scales should be culturally appropriate in each country, and the questionnaires should go through a thorough review process. Also keep in mind that as with any global project, the increased number of both internal team members and external contractors can have a major impact on the amount of time required for project management.

6.8 PERSONAS

Written by Lene Nielsen

The personas method has become widely used in many countries; for example, in Denmark many public Web sites are redesigned using the method, and in Japan the company Daishinsha has used the method for more than 60 projects since 2001. Personas are generally created for the different types of user groups within one country. Creating cross-cultural personas is not an easy task because the differences within a global user group are usually larger than those in just one country. This makes it challenging to create personas that are representative of the entire group.

This section reflects on the problems you may encounter when managing and executing a global project that uses personas. The process of creating personas is usually divided into different steps (Nielsen, 2007). An example of such a division is shown in Figure 6.11. We specifically address data collection, persona writing, scenario descriptions, and the acceptance of personas by involved partners.

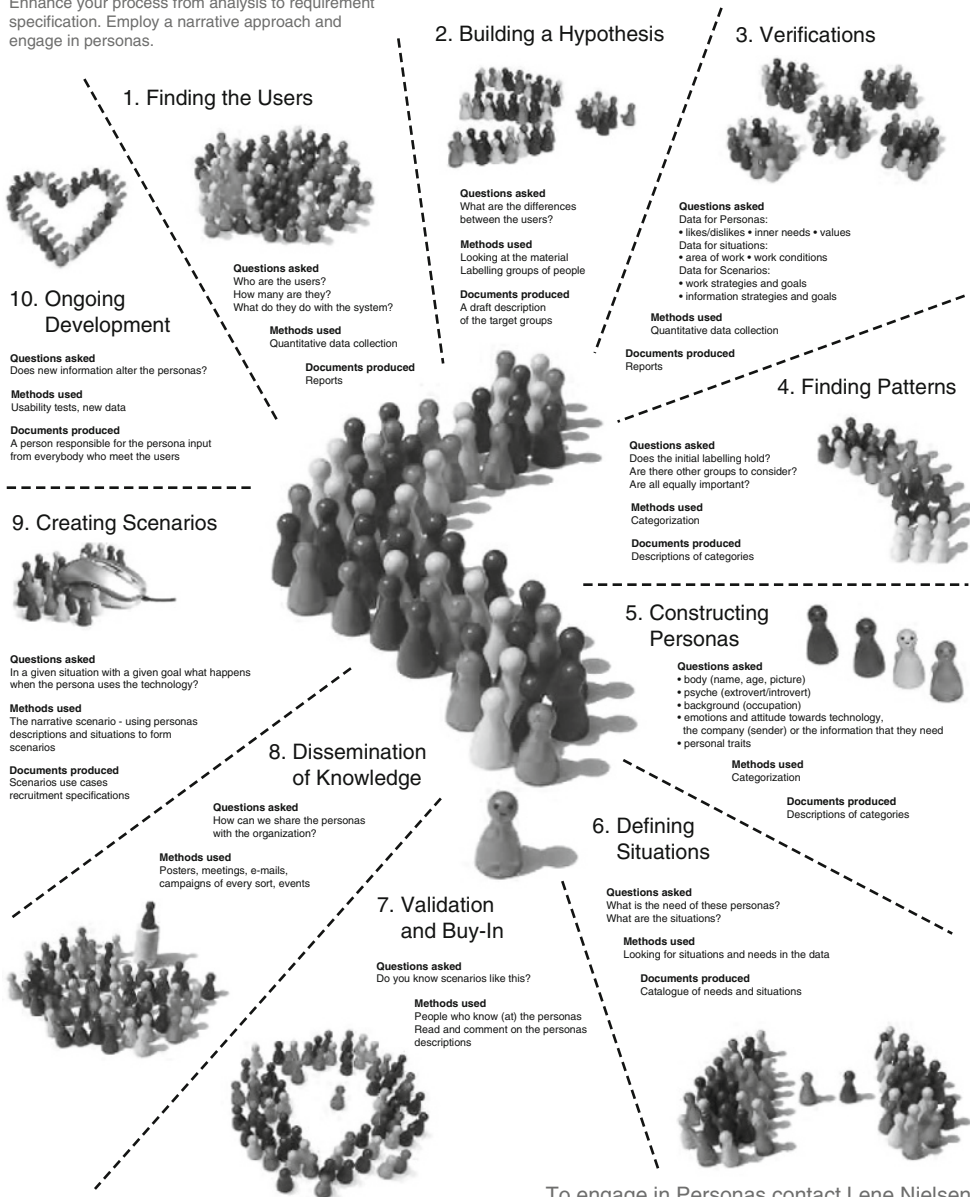
6.8.1 What is a persona?

A persona is a fictitious user constructed from different types of field data. Personas can originate from a combination of surveys, user interviews, observations, or any other user research method mentioned in this book. Personas are different from segmentation because data are collected with the focus on the problem area that the given project embraces. Personas are not just posters on a wall but, rather, are symbolic of key user groups and should be used to inform the ongoing design of a Web site or interface. For example, a designer can use a persona to imagine and visualize the end users' design preferences and needs. The personas must raise engagement in the reader in order to get the reader to be able to imagine the persona in a future use situation (Nielsen, 2004). The activities of the personas are typically explored in stories – scenarios – that describe future possibilities in an easily accessible way and in a format that is easy to change.

10 steps to Personas

Based on the method “Engaging Personas and Narrative Scenarios” (2004) by Ph.D. Lene Nielsen

Enhance your process from analysis to requirement specification. Employ a narrative approach and engage in personas.



© Snitker & Co. 2007 • Fotos by Sanjo Gjenero

To engage in Personas contact Lene Nielsen,
ln@snitker.com

■ FIGURE 6.11 Approach for creating personas in 10 steps (Nielsen, 2007).

Most often, the outline of the persona is a written description accompanied by a photo of the imagined user. The writing can have the form of a description, with a bulleted list that highlights certain criteria possessed by the user (e.g., age, sex, occupation, life situation, tools, resources, needs, and goals). It can also be posters portraying the persona in typical situations with fictive sentences describing the persona.

The reasons for using personas are many:

- The method maintains user focus and helps to focus on specific users.
- Personas are a way to make design assumptions explicit, to create informed design choices, and to investigate design ideas.
- Personas communicate data in an engaging way, can improve communication with stakeholders and designers, and encourage consensus.
- The method contributes to marketing and strategy.

The description is not just any kind of document; it is aimed at a known reader (the project team member), who has different needs than the persona that is portrayed. The personas are constructed to get the reader to identify with the personas and through this understanding to make informed decisions. The construction of the personas and the reading of the descriptions can always create misunderstandings, but this is even more likely when a project is cross-cultural.

6.8.2 The personas method in a cultural context

In our encounters with people we do not know – fictitious or in real life, belonging to our own nationality or another – we do not at first see the person as possessing a unique constellation of characteristics. Instead, we add the person to a previously formed category built on knowledge of meetings with others (Schutz, 1976). Later on, when an in-depth knowledge of the person is formed, the category is broken and the stereotype transforms into a personal character.

The personas method draws on this ability by creating personal characters that transcend the stereotypes. This makes it possible to understand people of different cultures and to imagine their life. In the following sections, we go through the personas process with the purpose of highlighting difficult aspects in dealing with cross-cultural personas projects.

6.8.3 Collect data

The initial step in the personas process is to gather data with respect to users' needs, goals, and attitudes toward the problem area. The outcome of this data collection phase is to understand how users differ in order to determine the final number of personas needed.

When planning a global project, the researcher must consider the differences within the local markets and decide on the number of countries from which to collect data. For example, a global bank wants to address its different customers. What is the starting point when deciding how many countries to get data from? How can we know which users are important to interview in each country? There are several ways of finding a starting point:

- Log file analysis might point to different and similar behavior among countries.
- Scientific reports are available on the Internet and can be used as a foundation for decision making. You can call researchers from the problem area, with the intention of getting a brief overview of research.
- You can use marketing segmentation studies available from the bank.
- Theories on differences between cultures can also be used to get an overview of differences (Hofstede, 1980).

Large global companies often have much information about their users (reports from marketing, call centers, etc.). Therefore, focusing on the problem area is even more important. Without focus, the amount of data can be overwhelming. The data can to some extent substitute real-life meetings with users, which is very practical for cross-cultural studies.

When analyzing the data in cross-cultural projects, it is vital to leave behind existing ideas of cultural differences. Analysis of data may show new patterns of people exhibiting similar attitudes across nations. Consider, for example, a global bank that wants to address its different customers. Examining trust in banks might show unexpected similarities across nations. For example, trust might be similar in Denmark and China, whereas it could be very different in Iceland.

The segmentation of users must be defined before researchers can begin to actually create a persona. In cross-cultural projects, this definition process might include discussing the observed differences between users with local partners to get their interpretation of the findings.

6.8.4 Write personas

When writing personas, it is not possible, nor is it necessary, to present all information about each persona. The reader may have to infer some missing information from expectations, knowledge of the depicted persona, and his or her own cultural background (Bordwell, 1997). Readers from different countries possibly have different expectations, knowledge, and backgrounds.

6.8.5 Create scenarios

It is in the scenarios that ideas of how the product is used arise, and it is the scenarios that test ideas of interaction from the personas' standpoint. Therefore, we argue that personas have to be combined with scenarios; otherwise, they lose much of their potential value.

A scenario is similar to a story; it has a main character (the persona), a setting (somewhere the action takes place), a goal (what the persona wants to achieve), actions that lead to the goal (interactions with the product), and obstacles that hinder the way to the goal. A scenario begins with a situation. For cross-cultural projects, there may be situations unique to a single country and therefore these situations should be considered in the scenarios.

Scenarios can be tricky for cross-cultural projects. Here, we deal with designing a future situation and imagining an unknown future. If a Westerner has to make a scenario for an Indian housewife, it might be difficult to know what her daily tasks are. Imagining how a product would fit into her daily life would be difficult, no matter how well it is described by the persona description, reports, photos, or even videos. Here it is essential to invite locals to participate in the scenario process to match the reality to the imagined future.

6.8.6 Ensure distribution and acceptance

To ensure personas are a part of a user-centered development process, the personas need to be distributed to all stakeholders within the project. Also, it is not only the personas that need to be distributed but also the data on which they are based – the foundation document (Pruitt & Adlin, 2006). This might be even more important for cross-cultural projects than for single-nation projects because it is vital that the stakeholders can get an easy overview of field data to understand the persona descriptions and to be able to track the arguments behind the descriptions.

An effective way of ensuring acceptance of the personas is to involve the project team members (and other stakeholders) in the persona development. Involving team members secures buy-in, understanding of the method, and dissemination of knowledge. Most often, it is by participating that many appreciate the strength of this method and understand how to use the personas as a tool to improve designs. This is also the case in cross-cultural projects. Therefore it is very important to find out who the project stakeholders are that should be involved. In cross-cultural projects, the list of stakeholders might be longer than that in local projects.

6.8.7 Summary

Cross-cultural projects put pressure on both the reader and the writer of the personas. If you are writing the personas, you have to understand the data and, as noted previously, use your own cultural background in the process of understanding. In the writing process, this can lead to misunderstandings and misconceptions.

The reader of the personas faces the same problems of having to use his or her own cultural background when reading the descriptions and trying to understand the personas.

To reduce misunderstanding, the best approach is to be sensitive about the cultural differences between your readers and to ensure that the descriptions are verified by locals. Their comments and insights might prevent misunderstandings and incoherent information.

Expressed differently, the traditional media model of “sender – message – receiver” is transformed in cross-cultural personas projects to a “sender – message – local moderation – receiver” model.

REFERENCES

- Bevan, N. (2008). *UX, usability and ISO standards*. Retrieved from http://www.cs.tut.fi/ihte/CHI08_workshop/papers/Bevan_UXEM_CHI08_06April08.pdf.
- Bordwell, D. (1997). *Narration in the fiction film*. London: Routledge.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Courage, C., & Baxter, K. (2005). *Understanding your users. A practical guide to user requirements, methods, tools, & techniques*. Boston: Morgan Kaufmann.
- Hofstede, G. (1980). *Culture's consequence: Comparing values, behaviors, institutions and organizations across nations*. London: Sage.

- iPerceptions. (2008). *4Q, an Avinash Kaushik and iPerceptions collaboration – free simple survey solution*. Retrieved from <http://www.4qsurvey.com>.
- Monster.com. (2009). *The world's leading career network*. Retrieved from <http://www.monster.es/geo/siteselection.asp>.
- National Institute of Standards and Technology. (2007). *Common industry specification for usability – requirements NISTIR 7432*. Retrieved from <http://zing.ncsl.nist.gov/iusr/documents/CISU-R-IR7432.pdf>.
- Nielsen, L. (2004). *Engaging personas and narrative scenarios*. PhD dissertation. Copenhagen: Samfundslitteratur.
- Nielsen, L. (2007). Ten steps to personas. *Journal of HCI Vistas*. Retrieved from <http://www.hceye.org/UsabilityInsights/?p = 73>.
- Pruitt, J., & Adlin, T. (2006). *The persona lifecycle: Keeping people in mind throughout product design*. Boston: Morgan Kaufmann.
- Rohrer, C. (2008). *When to use which user experience research method*. Retrieved from <http://www.useit.com/alertbox/user-research-methods.html>.
- Schutz, A. (1976). The stranger. In: *Collected papers – Studies in social theory* (pp. 91–103).
- Siegel, D. A., & Dray, S. M. (2005). Making the business case for international user centered design. In R. G. Bias & D. Mayhew (Eds.), *Cost-justifying usability: An update for the Internet age* (pp. 317–353). Boston: Morgan Kaufmann.
- UXalliance. (2008). Retrieved from <http://www.uxalliance.com>.
- Web Analytics Association. (2008). Retrieved from <http://www.webanalyticsassociation.org/aboutus/>.