Planning for usability testing

Whether you have the luxury of time or perhaps only a few days, you need to set aside time to plan for usability testing. The planning process can be divided into a number of steps, which begin with pulling together the people who will be part of the planning process and setting a time for the planning meeting. In the planning meeting, you and your team members will get as much done as you can in the time you have and then continue working on the test materials after the meeting and before testing.

This chapter provides essential test planning steps, including:

1. Scheduling the planning meeting, in which you:
   - establish test goals
   - determine how to test the product
   - agree on user subgroup(s)
   - determine participant incentive
   - draft the screener(s) for recruiting participants
   - create scenarios based on tasks that match test goals
   - determine quantitative and qualitative feedback methods
   - set dates for testing and deliverables

2. Writing the test plan
   - writing an informal test plan
   - writing a formal test plan
Scheduling the planning meeting

Planning a usability test begins with knowing who should be involved and what needs to be done. The process begins with scheduling the planning meeting and inviting the project’s stakeholders to attend. If everyone is co-located, you can meet in person. If all or part of the team is distributed, you can set up a meeting with GoToMeeting or another online meeting collaboration tool or via conference call.

Once you’ve determined who should be invited to the meeting, you need an agenda so that the attendees can prepare and bring the appropriate information to make the meeting productive. An agenda can be crafted from the generic one shown in Figure 5.1. This one shows you not only

---

### Planning Meeting Agenda

1. Establish test goals based on:
   a. How much money/budget you have for testing—this information affects many planning decisions, such as the size of the test (including whether this is a single study or part of an iterative testing cycle), the number of participants, the cost for recruiting participants, paying stipends to participants, and other costs, such as renting or allocating space, etc.
   b. Where the product is in development—this information affects the type of testing you will do.
   c. Who is sponsoring the test—this affects the focus of the study. Concerns of marketing will be different from concerns of technical support/user assistance, and so forth.

2. Determine how to test the product—this discussion may involve a review of the product and assessment of its status in development. The requirements for testing will be shaped not only by the status of the product being tested but also by the location of the participants: local, remote, or a combination.

3. Agree on the user subgroup(s)—depending on your budget for testing, this could mean settling on one subgroup or, with more money for testing, two or more subgroups.

4. Determine participant incentive—again, a budget issue, but this discussion often requires a review of options and agreement on what’s appropriate in the situation.

5. Draft the screener(s) for recruiting participants—this critical part of planning is tied to the user subgroups for this test.

6. Create scenarios—based on selecting tasks that match test goals.

7. Determine quantitative and qualitative feedback methods—tied to your goals for the test.

8. Set dates for the following:
   a. Testing
   b. Drafts of test materials for review (beginning with the screener)
   c. Final test materials (for test plan)
   d. Product ready for test setup (walkthrough)
   e. Post-test deliverables—agree on what these are, when they will be delivered, and in what format (written, oral presentation, video highlights, combination)

---

**Figure 5.1** This agenda for a planning meeting provides the topics with planning notes.
the items in a typical planning meeting but the rationale for including
them in the discussion.

Unless you can set aside a half day or more for the planning meeting,
you will probably not be able to produce all of the documents listed
on the agenda, so you need to have a plan for assigning tasks to team
members and sharing drafts of these documents for feedback and
approval after the meeting.

This agenda can be daunting, since there are a lot of decisions that
have to be made. That's why it's important to send out your agenda in
advance, identifying specific items that need to be prepared in advance
(and who should prepare them) and getting a block of time committed
so that you can make significant headway in the meeting. If you can get
only a few of the items on this list done in the meeting, make your top
priority getting everyone to agree on the test goals, user groups, and key
tasks. The outcome of a successful planning meeting is agreement on
these key items for the usability test and a list of follow-up actions with
due dates.

Establish test goals

As is frequently the case when you begin planning a usability test,
your team or sponsor wants to learn *everything* about the usability of
the product. As a result, the wish list can be quite lengthy. Yet the
practicalities of time and budget—not to mention the stamina of
participants in testing sessions—limit the scope of what you can do in
any particular test. So, how do you choose what to test?

You begin by setting your test goals. This is your first agenda item and
top priority for planning the test. Test goals focus on what you want
to learn about your users’ experience with the product at the point in
development where you will be testing. If this is your first usability test,
you need to decide what’s most important to learn from your users.
Maybe there’s an issue that the team has debated, and you’d like to get
user input to know which direction to take. Or you’ve gotten information
from customer support that a certain feature of the product is causing
problems for users, so you’d like to understand what the problem is from
observing users working with the feature. Or, if it’s a new product and
you want to know whether users understand what it is and how to use it,
you can focus on the new users’ experience to understand whether their
mental model for the product matches your design.
If you are planning for a follow-up study, you can set goals to learn whether earlier issues have been addressed in the redesign resulting from the prior study. Or you can set goals around new features you’ve added to the product.

If you’re not sure how to identify your testing goals, you could use criteria such as Whitney Quesenbery’s 5Es—Efficient, Effective, Engaging, Error tolerant, Easy to learn—to shape your discussion. Using the Es, your team can decide how to set goals for your study and how to measure whether these goals are met.

Here are some examples of goal setting using the 5Es:

- **Efficient**—Can users find the information they need to complete tasks without assistance? Can users perform a process within a predetermined timeframe?

- **Effective**—Can users successfully place an order or sign up for a service?

- **Engaging**—Do users rate their experience as satisfying or enjoyable? Do their comments (and body language) suggest that they are having a positive experience?

- **Error tolerant**—Do users experience errors? If so, how many? And when they experience errors, do they recover successfully? If they receive error messages, do they understand them?

- **Easy to learn**—Can users get started right away? Does their ability to do tasks improve as they become familiar with the system? Does the system architecture match their mental model for the way they expect the system to work?

These criteria and the relevant questions associated with each one not only shape the task list and the scenarios you will create, but also help you determine what you want to learn from observations and what you want to learn from post-task and post-test feedback methods.

Not included in this list of goals—but so important for understanding all of your users—is the goal of accessibility. Accessibility goals require special considerations and specialized recruiting for participants, so you may not be equipped to put this on your list for your first test. But, as you become experienced with the basic planning requirements for testing, you will likely want to understand the usability of your product for all users, and that means understanding the usability of your product.
Scheduling the planning meeting

for people with disabilities. The following sidebar gives you some useful information about accessibility goals and resources.

Accessibility goals are in everyone's interest

Is accessibility a goal? It certainly should be—if not for legal reasons, then for practical reasons. Knowing whether your product is accessible for people with disabilities or limitations imposed for other reasons—such as diminished eyesight or mobility brought on by age or infirmity—helps you understand how to reach a part of your user population that you may not be currently reaching. Yet accessibility is often viewed as a nice-to-have goal if time and resources allow.

Legal reasons to address accessibility

The Web Content Accessibility Guidelines (WCAG) are a recommendation of the World Wide Web Consortium (W3C). W3C is a web standards organization that launched the Web Accessibility Initiative (WAI) in 1997 to ensure that W3C guidelines support access for all people. Europe and part of Asia have similar initiatives.

In the United States, Section 508 (a 1998 amendment to the Rehabilitation Act of 1973) sets standards for accessibility of information technology. It applies to all electronic and information technology products that are developed, purchased, or maintained by the federal government, which gives it a broad reach into private industry. State governments often use 508 as the baseline for their own state standards.

The Americans with Disabilities Act (ADA) affects private business as well, extending protection to all people with disabilities to provide them with equal access.

What does this alphabet soup of rules, regulations, and organizations mean for your product? It means there is plenty of information available on how and why to address accessibility, with lots of help on how to do it. These guidelines codify the requirements into a series of rules, but there’s no substitute for knowing your users.
Making the business case for testing for accessibility

There’s an excellent business case to be made for addressing accessibility. When improvements are made for people with disabilities, studies have shown that the user experience also improves for people without disabilities.

Accessibility benefits people without disabilities, including:

- older people
- people with low literacy
- people without native language fluency
- people with low bandwidth connections or older technology
- people with low web literacy skills

In other words, all boats rise when you address accessibility.

Some major corporations, including Microsoft, IBM, and Fujitsu, have championed accessibility, publishing guidelines for universal access. Fidelity.com has made accessibility a core part of its planning and testing methodology.

Others have been made aware of the need to address accessibility the hard way. Take Target as a high-profile example. The big-box discount retailer assumed that the Americans with Disabilities Act did not extend to websites, particularly when the company provides access through its brick-and-mortar stores. That all changed when the National Federation of the Blind brought suit against Target and won. Similar lawsuits or structured settlements have been settled against hotels, travel sites, and banks.

If you find yourself needing to make the business case for addressing accessibility, a number of resources on the Internet can give you ammunition.
Determine how to test the product

This agenda item has several parts. In determining how to test the product, you need to decide:

- **What to test**—based on where the product is in development
- **Where to conduct the test**—based on choices for lab testing, field testing, testing remotely, some combination, or another option
- **How to test**—based on resources, timing, and your goals

**What to test—Product**

Your discussion about what to test will focus on where the product will be in development when you want to conduct the usability test.

Ideally, you will be testing iteratively throughout product development. If, however, this is your company’s first usability test, it is likely that management or the test sponsor will want to wait to test the product until it’s nearly complete. This type of testing is called *summative* evaluation because it assesses the usability of the product at the end of development. It’s useful and valuable to do summative testing when you want to confirm that requirements have been met for the product.

But if this is your first test, you should work to persuade the decision makers to test earlier in development so that the findings from testing can lead to more user-centered development of the product. This type of testing is called *formative* evaluation, and it can be used to test very early paper prototypes or partially developed products, the information architecture of the product, or a particular feature by itself.

**Where to test—Location**

Once you have decided where the product will be in development when you do the usability test and what your goals are for the test, you need to decide where you will conduct the test:

- *In a lab*—by testing in your own lab or one that you will rent
- *In a conference room*—by reserving a room for the test
- *In the field*—by going to the users in their environment
- *At a distance*—by testing remotely
If you have a lab, you’ll probably want to use it, especially if this is your first test of the product. Of course, you don’t need a lab to conduct testing. Reserving a conference room or renting a local facility works just fine.

If you’ve done lab testing of the product already, you may want to get out into the field so that you can see how the product works in the users’ environment. A combination of lab testing in one study and field testing in another study gives you a fuller picture of your users’ experience because it combines controlled and natural testing environments.

You can also do remote testing, which gives you the reach to your users wherever they are and in their own environments. You can decide that you want to use remote testing for all of your test sessions, or you can decide to combine some remote and some lab testing in a single study.

How to test—Design

As you’ve moved through your agenda, you’ve made some important decisions about the goals for your usability test and the status of the product you’ll be testing. Now you need to discuss the type of test design you would like to set up. Some methods for structuring the test design include these:

- “Typical” test of the product—when you present users with a number of tasks within scenarios, which gives you similar feedback on their experience with your product. Usually formative for products in development.

- Benchmarking—when you test your product with users to establish metrics, or benchmarks, for the product, as well as requirements for new product development. Usually summative for completed products.

- Comparison of designs—when you present users with two or more designs so that you can see whether a preference emerges.

- Competitive evaluation—when you present users with tasks to complete in your product and one or more competitor products to learn their preferences or to measure your product against the competition.
The decision on the type of test design you want affects the scenarios you will create and the post-task and post-test feedback mechanisms you will use. In addition to making this decision, you will also need to decide how much time you want for each test session. If you are using the “typical” approach, you will probably set up scenarios that can fit into sessions of an hour to 1½ hours. Longer sessions than this can tax the concentration of participants and the team, so if you need to test in a longer session, you probably want to build in a break in the middle.

Using the one-hour session as an example, a typical test day will have setup time of an hour, followed by participant sessions of an hour each, with short breaks in between and time for lunch in the middle. In a day, you will likely get to see five or six participants. If you run your findings meeting at the end of the day, you can add a couple of hours to the length of the day. This test schedule makes for a long but productive day.

But what happens if you don’t have even a day for testing? What if you don’t have the budget for testing in the usual way? What if your development cycle is so tight that the team can’t wait for results before moving on? All of these questions are realistic and frequently asked. The answers come in the form of some fast and effective testing methods that speed up the process of testing and delivering results. In the world of ever-tightening schedules—made more so for those companies using an agile development methodology—these faster testing techniques keep usability in the picture. The following sidebar gives you some of these techniques.

---

**Testing faster, cheaper**

Agile programming requires agile usability testing methods

In a waterfall development methodology, each phase of development is clearly defined on a timeline with milestones to mark progress in distinct phases. Planning for usability testing is a matter of scheduling testing at specific points along the development timeline.
Nowadays, a growing number of companies have switched to the agile development process, which is making it more challenging to insert usability testing.

If you are not familiar with the agile method, here’s how it works in a nutshell: Design teams work in very short development cycles, called sprints, of one week to one month, typically several weeks. In each sprint, the goal is to get a feature or a group of features designed and coded. The overall goal is to deliver working software early and frequently.

How does usability testing fit into this pace of activity? As some user experience teams have found, the solution lies in the adoption of some fast and agile testing methods. Building user experience methods into the agile development process works particularly well when the user experience team works in parallel with the development team so that UCD practices can be a separate but coordinated part of product development.

The RITE method is well suited to agile and other rapid development processes

RITE stands for Rapid Iterative Testing and Evaluation, a great name that says it all. Developed by the user testing team at Microsoft’s Games Studios, RITE addresses the business need to make fast changes to a design as soon as a problem is identified. RITE is quick, and it is agile. Using the RITE method, you can schedule a few participants during a sprint or just after it and fix whatever problems you see before having a few more users work with the product to confirm that the fixes work.

Making RITE work requires full team commitment, including key decision makers who are both knowledgeable of the product design and able to approve changes. The methodology works like this:

- Key decision makers observe the participants in testing.
- The findings are analyzed immediately after a test session concludes or, in some cases, at the end of the day.
• As soon as there is agreement that a problem exists and the solution is known, the change is made immediately. If the team isn’t sure they’re seeing a real problem, testing proceeds until the problem and solution become clearer.

• The changed interface is retested with the next participant.

Unlike traditional usability testing—in which you wait until you have seen all of the users, then analyze what you saw, then recommend changes, then make the changes—the RITE method focuses on redesigning to fix problems and then confirming that the solution works with users.

Other testing techniques provide fast feedback

In addition to the RITE method, other techniques may suit the needs of your product, your schedule, and your goals. Among the growing number that are being used are these:

• **Weekly testing**—Participants are recruited and scheduled every week so you can test whatever you have ready. Since participant recruiting has the longest lead time in planning for testing, this technique takes the delay out of the picture. If your company has several development teams working on different products, this weekly testing schedule can be managed with a sign-up sheet for testing on a first-come, first-served basis. If you don’t have an internal usability group to do this recruiting and scheduling for you, you can set it up yourself.

• **Quick testing at a conference**—You can plan for testing at a conference or trade show, which gives you access to lots of potential users. This works particularly well when your users are hard to recruit. Using your company’s exhibit booth to catch interested participants for a short, informal usability session, you can get a fast response to new features and functions that you are trying out in development.

• **Five-second tests**—This technique can be done with a paper prototype at a place where your users gather, such as a shopping

For more on using RITE and integrating best usability practices in agile development cycles, see the first-quarter issue of UPA’s *User Experience Magazine*, 2010, which has several articles on this topic.

mall or park, or you can use a free web tool available at www.fivesecondtest.com. Using the web tool, you upload a screen you want users to review for five seconds, then take some action. With this technique, you get feedback on first impressions and other aspects of your design, such as where users first click and what they remember about your website after only five seconds.

The point about all of these techniques is that there are many clever options to choose from, no matter how little time or budget you have. In your planning meeting, you can keep the focus on choosing the right method, knowing that there will be one that best fits your situation.

Agree on user subgroups

At this point in your planning, you have determined the goals for your study and the type of test you will conduct. You now need to agree on the user subgroups for your study.

As you know, you can conduct a study with 5 users and get excellent results as long as the users are all from the same subgroup. If you have time and budget to test with 10 participants, you can identify two subgroups, or possibly even three. The more participants you plan to recruit, the more subgroups you can draw from.

Whether you have budget or time for only one day of testing or for several days, you need to decide on the subgroup or subgroups you want represented in your study. This decision can be difficult to make, because your team or your sponsor often wants to know about the user experience for many different groups of users. In your planning meeting, you need to get buy-in on who your users will be for this study.

If you have developed personas, this makes the task a bit easier because it gives you a starting point to discuss which personas you want to include. However, a persona represents a type of user, and there could be a number of subgroups within a single persona. Also, personas do not typically include the specific characteristics you will be seeking to match for your study.
So, you need to come up with a list of characteristics for each subgroup. Let’s call this list of characteristics for a particular subgroup a user profile. You will need a separate user profile for each subgroup. In your planning meeting, you may have time to generate only the list of characteristics for each subgroup. After the meeting, you can convert these subgroup characteristics into the screener, then circulate it for feedback and approval by the team and any other stakeholders before recruiting begins. If you need to start recruiting right away, this is the first deliverable you need to finalize during or after the meeting.

Define the characteristics of a subgroup

A single, definitive list of characteristics for each subgroup would not be possible to create, but here are some characteristics that typically generate differences among subgroups, using the examples of software and websites/web applications.

For software:

- familiarity with the type of product you are testing
- familiarity with your product—current or earlier version
- domain knowledge as it relates to your product
- technical skills as they relate to use of your product
- computer skills
  - computer usage
  - device usage
  - Internet usage
- software skills
  - applications
  - usage
- job category
  - job title and type of work relevant to your product
  - could include other categories such as:
    - student
    - retiree
    - stay-at-home parent
For websites and web applications:

- familiarity with the web, based on types of usage/activities and amount of time per week/month
- familiarity with websites/applications that are competitors or that share the same space as your website or application

You notice that I did not organize the characteristics by “novice” versus “expert.” These terms are extremely difficult to define. And asking people to categorize themselves rarely works, since they will interpret the meaning of the terms in widely different and generally inconsistent ways.

To better place users in categories of expertise, it helps to focus on their experience with the tasks or tools you will use during testing. Once you’ve established some minimum and maximum ranges, you can group potential participants into experts, intermediates, or novices. You can then decide whether you want to see users in all three of these categories or only one or two.

For example, one study may focus on new user experience with novices. Another may focus on new user experience with people who have used relevant, related products, but not your product. Another study may focus on experienced users of your product who are being introduced to a design change or new features. Or, if you can recruit from several subgroups representing these users, you can combine them in your study.

Focus on user motivation

For all subgroups, matching user motivation to study goals is the most important factor in deciding who to recruit. For instance, if you are testing a website that provides information about cars to allow users to make a purchasing decision, all subgroups of users must share the common goal of being interested in this information because they are planning a new car purchase. Perhaps you set a limit of intent to purchase in the next six months. Without this real motivation, participants in the study are likely to treat the tasks as exercises that have no real meaning for them.

Mix some characteristics within a subgroup

Now that you’ve established the subgroups of users you want to use, you can mix in a number of characteristics within a subgroup while still
maintaining consistency among the critical factors you’ve identified regarding motivation, skill level, and experience. Depending on the goals of your study and the variety of users within subgroups, you can get a healthy mix from the following characteristics:

- **Age**—The range can cover the entire user population, or you can set a smaller range, such as 18- to 30-year-olds.
- **Gender**—Typically, you want a 50/50 mix or close to it, unless your users are mostly (or all) male or female.
- **Education**—If all levels of education are represented in your users, the range can be from a high school education (or less) to a Ph.D. However, you may want to narrow the range. For instance, if you are seeking only low-literacy users, then defining the education level is a critical factor (although there can be other reasons for low literacy).
- **Language**—Sometimes it’s appropriate to have second-language users, which does not necessarily mean low-literacy users.
- **Ethnicity**—You may want to have a mix of ethnicities if the product’s users include these ethnicities. If there is a dominant ethnicity, this becomes a critical factor.
- **Disabilities**—People with disabilities can be a subgroup of its own, which would move it to a separate category, or you may want to include people with disabilities. This can be determined by either seeking people with disabilities or not ruling out people with disabilities if you identify them in screening.
- **Economic factors**—Household or individual income can be helpful in understanding purchasing power, market differentiation, and other factors. If a minimum economic requirement is established, you can solicit information by learning whether respondents:
  - own or rent a residence
  - own or lease a car or truck—one or more, brand, model, year
  - own products like the one you’ll be testing—how many, what brands

Or if specific income brackets are required for your study, a question about household or individual income can be used.

**Combine characteristics in a user profile**

It won’t be necessary to include all of the characteristics you listed in your user profiles. But it will be necessary to decide what’s most
relevant to your study, based on your goals. Some examples from
different types of studies will show you how to choose what to include
and what to leave out.

**Example: File Transfer Protocol (FTP) software**

Two subgroups were identified by the team as novices and advanced
users on the basis of these characteristics.

**Novice user:**

- complete beginner in the domain—no prior experience or even concept for FTP
- must have home computer with network connection
- minimum one year Internet experience
- minimum ½ hour/day Internet use (or four hours per week)
  - surf Internet for information, shopping
  - download information from Internet
  - use e-mail, use attachments
- optional technical abilities can include use of digital camera, or MP3s or downloaded music
- no network concept understanding (terms such as host and IP address are likely to have no meaning)
- must not have had experience designing or building a website (except, perhaps, using a wizard to create a web space at a portal site such as Yahoo!)
- must have interest in creating a website or using a host site to share or post files (for example, eBay)
- must not have had experience loading files to an FTP server
- must be native/fluently English speakers (no English-as-a-second-language [ESL] speakers)
- age not an issue—no restrictions, no need for age distribution

**Advanced user** (based on team-generated persona for technical and power Internet user):

- job title/description: Webmaster, IT manager, or IT professional
• goal: seeking to gain productivity
• current FTP user, but not a current or recent customer (may select users with prior experience with software, earlier version)
• FTP experience with one or more other FTP products:
  ◦ list of products here
  ◦ other (for any product not included in the list)

Questions to ask potential participants:
• How long and how often have you used FTP products (for each of the preceding)?
• How would you make a second simultaneous connection?
• Do you manage encrypted files? If so, how?

Example: Website for information about distance learning courses and programs throughout a public state university system
• ten participants in two groups
  ◦ current undergraduate students
    – traditional ages 18–21
    – nontraditional ages 22–35
  ◦ graduate students
    – current
    – prospective
• must have interest or experience in taking distance learning (online) course or program
• must not have visited the website being studied
• priority for those who are interested in or currently seeking an education degree, especially in math and science
• must use computer a minimum of three hours per week (not including e-mail) for a variety of things, which could include school-related research, other information seeking, shopping, social networking, bill paying
• equal number of men and women
• some second-language speakers
• diverse ethnicity
Example: Self-installation of a digital cable box

- must have cable/satellite TV now or indicate that they are thinking of getting it
- must indicate preference (among choices) for attempting the self-installation versus paying a fee for a professional installation
- for those who are current cable subscribers:
  - must not have done the cable installation for their current TV themselves
  - must have some other equipment connected to the TV, such as VCR, DVD, stereo
- must have some comfort level_experience using or performing at least one of the following:
  - using computer/Internet—at least three hours per week
  - installing hardware—printer or fax—to their computer
  - adding or upgrading a component to a computer, such as a memory card
- equal number of men and women
- variety of
  - ages between 18 and 55
  - household incomes
  - ethnicities
  - education levels from less than high school diploma to college degree
- seek some ESL speakers, especially Spanish as first-language speakers

Determine participant incentive

Because you are asking people to take time out of their busy lives to help you understand their experience with your product, you need to compensate them for their effort. This is a tricky subject, though, because you want to make the incentive feel like a thank-you gift and not a bribe. In other words, you don’t want the incentive to influence the remarks they will make as they work with the product and complete questionnaires. You also want to avoid recruiting “career participants” who do this sort of thing strictly for the incentive.
Scheduling the planning meeting

What’s an appropriate incentive?

When it comes to determining the appropriate incentive, as with most things associated with usability testing, “It depends.” The amount will vary according to the general cost of living for your city and the convenience of getting to your testing location. The amount could be US$50–60, particularly if it’s easy to get to your testing location. In Atlanta, with our awful traffic congestion, we find it difficult to recruit for less than US$75 in participant incentive. This dollar amount is based on a typical study of about an hour, but it factors in the commute time as well as our request that participants check in 15 minutes early to complete paperwork.

If you are going to recruit from more than one subgroup of your user population, the incentive may not be the same for every subgroup. For instance, some of your users may be nonprofessionals, such as students, who can be recruited for less incentive than professionals in high-income salary ranges. Or you may be recruiting from a large pool of potential participants who are readily available. In this case, the incentive can be less than when you are recruiting hard-to-find or hard-to-get participants. For these hard-to-get participants, such as IT network managers, you will need to increase the incentive to US$150–200 or more.

Besides cash, common incentives include gift cards or debit cards. The face value of the card would match the appropriate cash incentive. It’s best to make the gift card or debit card generic enough to allow the participant to use it in many places or for many things. Amazon.com cards work for this, as one example. Another approach is to offer a variety of product-specific gift card options from which participants can choose. This works well when you are planning for a number of studies, because you can buy the various cards in advance and have them on hand.

What do you do when you don’t have money for incentives?

The participant incentive is part of the cost of doing a usability study, so the amount you can offer participants may be determined by the budget for testing. If your budget does not include money for incentives, there are ways to recruit without providing an incentive. Friends and family are one source for potential participants who will not likely need an incentive. If you find yourself with no choice but to recruit friends and family, you want to strive to get as close as possible to your real users in the subgroups you are interested in for this study.

Carolyn Snyder, a usability consultant, offers this advice when you’re responsible for getting the participants’ incentives from an ATM: “Make the amount match the way in which ATMs provide cash.” Since U.S. ATMs don’t typically dispense $5 bills, Carolyn suggests making the incentive some figure in multiples of $20s or $50s.
If your target participants are internal users, you may not have to provide an incentive. In this case, refreshments can be an appropriate incentive. If a small incentive is appropriate, it can be something like movie passes or a meal ticket for the company cafeteria or a coffee card from the local coffee shop.

In some cases, such as employees at the U.S. federal level and at some state agencies, participants cannot accept an incentive. When participants cannot accept an incentive, they may still have sufficient motivation to want to participate in the study.

Whatever incentive or motivation you can use to recruit participants, you need to decide what it is so that you can include it in the screener.

**Draft the screener for recruiting participants**

As you move through your agenda, the hard work of determining your subgroup or subgroups of users is now done. If this is all you get done in the planning meeting, that’s a significant part for an effective study.

You still need to draft the *screener*, which is the document that will be used to recruit participants.

Unless you have set aside a big block of time for the meeting, you probably won’t get the screener drafted during the meeting. Instead, you may need a separate meeting for this, or agreement for someone to take the lead on drafting and circulating the screener(s) for feedback and approval from interested team members. Whether you do the recruiting yourself or use a recruiting agency, you will want to get buy-in on the specifics of the screener so that whoever does the screening is working from an approved document.

If you’ve never drafted a screener before, you will benefit from getting feedback from everyone. Once you get comfortable with the process, you can ask those with an interest to opt in to the review process.

Figure 5.2 is an example of a screener for the usability study of the digital cable box self-installation, which you can use to help you prepare your screener.
Digital Cable Box Self-Installation Study

Study Dates: May 12–14, 20xx
Study Times: 8:30 AM to 5:30 PM
Recruiting Goal: 21 participants, plus backups

Special Notes for Recruiters:
- Equal gender distribution
- Some English-as-a-second-language speakers, preference for Spanish
- Representative age distribution (18 to 55 yrs)
- Representative education mix of no college / some college / college degree

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Recruited by:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name:</td>
<td></td>
</tr>
</tbody>
</table>

(Although we are seeking some candidates with English as a second language, all respondents must be able to communicate clearly in English. If you have doubt as to their ability to communicate clearly, terminate.)

Candidate speaks in clear, understandable voice?  
☐ Yes (continue)  ☐ No (Terminate)

Recruiter Introduction
Hello, my name is ____________________ from The Usability Center at Southern Polytechnic State University in Marietta, Georgia. We are conducting a study about how people learn to connect home entertainment products to their cable service, and would like to ask you a few questions. If you qualify for this study, we would like you to participate in a single one-hour session at our facility. You would be compensated $75.00 for your time.

If selected, will you allow us to videotape you?  
☐ Yes (continue)  ☐ No (Terminate)

Screener Questions
Because this study involves connecting home entertainment products to cable, we’re looking for people who own a TV and perhaps also some other equipment, such as a DVD, VCR, or stereo. Perhaps you already have cable or satellite service to support your use of this equipment. Or, you’re thinking of getting service in the near future.

Does this sound like you?  
☐ Yes (continue)  ☐ No (Terminate)

What type of television(s) do you have? (brand, size, etc.) (If none, terminate)

______________________________

Are any of your televisions connected to any recording or playback equipment, such as a DVD, VCR, or stereo? If so, what is the equipment? (brand, type, function, etc.)

______________________________

For your TV reception, what type of service do you currently have? Who is the provider? (Cable/Comcast, satellite/DISH Network). Other? (If none, terminate)

______________________________

Figure 5.2 This screener is for the digital cable box self-installation study.
Who installed it?
Company installation _______ Friend or relative _______  
Hired a consultant/contractor _______ Self-install _______ (Terminate)

Before answering the next question, please consider the following statement: It costs $41.95 for a professional installation of a digital entertainment package. If you do it yourself, the same package is only $9.95. However, if you are not able to perform the installation successfully and the vendor has to send a technician to assist, you may be charged an additional $35.00 for the help provided. Now, given the choice between a professional installation and a self-install, which would you choose?

☐ I would choose to do it myself for $9.95 because I am confident about my chances for success
— OR —
☐ I would choose to do it myself for $9.95 and I would be OK with paying $35.00 for help if I needed it
— OR —
☐ I would not choose to do it myself and would pay $41.95 for a professional installation. (Terminate)

Questions About Computer Use and Experience
Do you own a computer?  Yes ☐ No ☐
If no, do you use a computer?  If yes, ask where: (If they do not use a computer, terminate)
☐ At work  ☐ Library
☐ School  ☐ Other

What do you typically use the computer for?

When thinking about the time you spend using a computer in an average week, how many hours do you spend using it?
Would you say you spend: (Read list)

<table>
<thead>
<tr>
<th>Hours/Categories</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than three hours per week</td>
<td>1</td>
</tr>
<tr>
<td>Between three and five hours per week</td>
<td>2</td>
</tr>
<tr>
<td>Between six and nine hours per week</td>
<td>3</td>
</tr>
<tr>
<td>Between 10 and 20 hours per week</td>
<td>4</td>
</tr>
<tr>
<td>More than 20 hours per week</td>
<td>5</td>
</tr>
</tbody>
</table>

Have you **personally** ever installed an add-on or upgrade component to a computer, like memory or a graphics card?  ☐ Yes  ☐ No

(If yes) What was it?

Have you **ever** installed hardware, such as a printer or a fax to a computer?  ☐ Yes  ☐ No

(If yes) What was it?

Demographic/Other Questions:
Record gender from voice:  

<table>
<thead>
<tr>
<th>Gender</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>1</td>
</tr>
<tr>
<td>Female</td>
<td>2</td>
</tr>
</tbody>
</table>

Figure 5.2 This screener is for the digital cable box self-installation study. (Continued)
We want to recruit candidates from a **variety of age groups**. Please tell me in which **category your age is**.

- Less than 18 years (Terminate)
- 18–23 years
- 24–29 years
- 30–35 years
- 36–39 years
- 40–45 years
- 46–50 years
- 50–55 years
- 56 or older (Terminate)

Which of the following categories best describes your race or ethnic background? (Read list)

- American Indian or Alaskan Native
- Asian
- Black or African American (not of Hispanic origin)
- Hispanic or Latino
- Native Hawaiian or other Pacific Islander
- White or Caucasian (not of Hispanic origin)
- Other

Do you work?  Yes  No

(If yes) What type of work do you do?

And is that full-time or part-time? (Circle one)

What is your highest education level?

- Completed high school
- Some college
- Completed college. Degree in?

Which of the following categories includes your **household income**? (Read list)

- Under $50,000
- $50,000 to $75,000
- $75,001 to $100,000
- $100,001 to $125,000
- $125,001 to $150,000
- More than $150,000
- Prefer not to say

Thank you! This concludes our questionnaire.

If you are selected for this usability study, you will receive $75.00 for your participation. When you come to our Usability Center 15 minutes before your session starts, we will ask you to sign a release form that allows us to videotape your activity for research purposes.

If you are selected for this study, what's the best way to reach you to schedule your session?

- Cell: __________________________
- E-mail: _________________________
- Daytime number: _______________________
- Evening number: _______________________

**To be completed later:**

- Scheduled session day: _______________________
- Scheduled session time: _______________________

**Figure 5.2 (Continued)**
Create scenarios based on tasks that match test goals

The next item on your agenda is determining the tasks you want your users to do with the product. Now that you know who the users are for this study, you can match the tasks to their goals for the product and your goals for the study.

To decide on the tasks, which will be crafted into scenarios, think about the questions you want your users to answer. Figure 5.3 shows you how to move from a question about navigation on a website to a task to a scenario.

Create “real” scenarios

Scenarios need to feel real to all of your participants. Here are some of the elements that you need to address in writing realistic scenarios:

- Use the language of the user, not the product.
- Put the tasks into a context of use that matches the user’s world.
- Give the user a goal, not a list of steps to accomplish the task and reach the goal.
- Say as little as possible to present the goal. You don’t want to write a short story or overload the user with unnecessary details. And you don’t want to give away more information than the user would be expected to have to perform the task.
- In situations for which the system requires personal information from a user, provide this type of information to reduce unnecessary exposure of a user’s personal details. You may need to create a unique set of data for each participant, starting with a specific user name and password, which the system you are testing will recognize.
- In some cases, you may want users to use their own information to make the tasks more realistic and meaningful. In these situations, you need to review this requirement during screening to make sure participants will be comfortable and can bring the information they need. Be prepared to have a fake set of data available for use if some change their mind about using their actual personal information.
To support scenarios that ask users to respond to certain situations, provide a description to help users, such as:

- Information to enter into a text box or to put in a text message. For example, you can tell them that when prompted for the reason for their inquiry, they will respond that they are reporting a service outage. Or you can tell them that they need to send a text message to a friend to say that they’re going to be 15 minutes late.
Any other information to standardize users’ responses in a scenario. For example, you can specify how many of something to purchase, for whom, and how much to spend (with the credit card number you provide).

**Decide on your first scenario**

How do you want your users to first experience your product? What’s the best starting point? Deciding on your first scenario should be based on the answers to these questions, which are tied to your goals for the study. Specific questions to help shape the first scenario will also be dependent on the type of product you are testing.

For a website:

- Are you interested in how they find it?
- Or do you want them to start from the homepage?
- If they are current visitors, do you want them to start with a new feature or service?

For software:

- How does the user get started?
  - Is there an application they should install?
  - Or do you want to install the application and have them start by clicking on the application?
- If you’re focusing on a new feature, will they already know the basic features?
- If you’re focusing on instructions or tutorials, do you want them to
  - Start by taking the guided tour to learn the software?
  - Or start with a specific task to see what instructions or help they need?

For hardware:

- Are you interested in the out-of-box experience?
- Are you focusing on the documentation for setting up the product?
- Are you interested in learning what users can do without documentation?
Your first scenario is often designed to capture initial impressions

At whatever point you start the test, it’s often a goal to capture users’ first impressions, especially when the study is formative. A typical first scenario that I frequently recommend is called the “look and feel” scenario because it asks participants to look around the homepage or the first screen or page view of the product and share first impressions by responding to these sorts of prompts:

- What type of site or application is this?
- What do you think you can do here?
- Look over the tabs or links and share what you think these mean and what you think will happen if you click on them.
- What would be the first action you would take?
- Are there any words or labels that you don’t understand?
- What’s your general impression of the site from the homepage or first screen?

Your first scenario should be short

Planning a short first scenario has several advantages:

- It allows participants to see how this process is going to work and, if they have any initial anxiety about it, to get comfortable.
- It allows you to correct any unforeseen technical problems and other logistical issues.
- It gives you a natural stopping place to reinforce the importance of having the participant think out loud. Having this natural break to remind the participant about this is especially important if the moderator is not sitting with the participant.

Other scenarios come from your task list

After deciding how you want to start the test, you next need to decide how many other scenarios you want to create, using the technique that was shown in Figure 5.3. You probably have more tasks (and questions) than you have time, so you will need to set priorities for the scenarios you create and assign some estimates of the time for each scenario. You can also create some optional scenarios to use if there’s time remaining in some sessions.
Your scenarios could be organized in a number of ways, such as the ones that follow:

- a sequence of tasks that has to be performed in a particular order
- most frequently performed tasks
- tasks getting the most calls to the help desk
- the most critical tasks (which may not be tasks that are performed most frequently)
- new tasks/new features
- tasks or task flows that are the subject of internal debate (the user can show the team what works, what doesn’t)
- comparative tasks using alternate designs of your product or your product and a competitor’s product

Comparative/competitive testing requires special considerations

Comparative testing lets you try out different designs to learn user preferences. Competitive testing lets you assess your product against a competitor’s product or assess competitor products against each other. These types of tests can produce valuable insights in both formative and summative testing, but they require special handling to set up scenarios that provide balance and fairness.

If you have never done a usability test before, I wouldn’t recommend that you start with this type of testing. I also wouldn’t recommend testing more than two options until you get comfortable with handling the data, because the data analysis gets more complex with each additional product you include.

However, if you’ve done some basic testing before and the team’s goal for this study is to compare alternate versions of a design or certain tasks in your product with similar tasks in a competitor’s product, or to do a competitive evaluation of products currently on the market, then you will want to set up this type of test.
Here are some of the questions you will need to discuss. The answers will shape the way you set up your scenarios and design your study.

- **Will the same users test both products?** This type of testing is called *within-subjects* design. With the same participants using both products, you want to set up an A–B, B–A comparison. That means that half the participants will begin with the A product and half with the B product. Within the tasks, you may want to vary the order again, such that half the participants are exposed to task A first, with the other half exposed to task B first. This assumes that the tasks do not need to be performed in sequence. The within-subjects approach takes more time, clearly, because participants are performing the same tasks twice. However, the same participant tests both products.

- **Will half the users test one product and the other half test the competitor product or alternate design?** This type of testing is called *between-subjects* design. Because you have reduced the number of participants by half for testing each product, you will typically need more participants to get a clear picture of your users' preferences. However, the testing session takes less time than the within-subjects design because participants are working with only one product. Using this approach means that you have to screen participants very carefully to match the characteristics of the users for either product.

As you can see, there are advantages and disadvantages to each of these approaches. So, the team needs to decide which approach is a better match for the study's goals, time, and budget for testing. Once the approach is chosen, you can craft the scenarios to fit the situation.

**Decide how scenarios will end**

Your participants need to know when they have completed a task and you need to know when they think they are done. To get this feedback, you will want to write the ending into your scenarios. Typically, this includes a request that they tell you when they are done. An example is
something like this: “When you have completed this transaction, let us know that you are done.”

It's important to hear this from the participant, since it is a common finding in usability studies that the participant thinks the task is complete when, in fact, it is not. It is also sometimes the case that you observe the participant complete the task, but then you see that he or she continues beyond the completion point to confirm that it is done correctly.

Knowing when the participant thinks the task is finished can be critical to understanding whether users can complete a transaction successfully on their own and whether they feel confident that they have done it correctly. If you hear participants say, “I think I did that right,” you very likely have a problem you need to address.

**Decide how to provide the scenarios to the participant**

How will you distribute the scenarios to the participants? Consider these two approaches:

- Give the participant the first scenario. Ask the participant to tell you when he or she is done so that you can then give the participant the next scenario, or the post-task questionnaire and then the next scenario. Do this for each scenario.
- Give the participant all of the scenarios, and tell the participant to continue with the next scenario when he or she has completed each one.

The advantage to the one-at-a-time approach for providing scenarios is that you can interact with the participant between scenarios and you can control the number and order of the scenarios you give the participant. This method works well if you are on a tight schedule or are varying the order of the scenarios for different users.

The advantage to providing all of the scenarios at once is efficiency, in that the participant can move through the scenarios, along with any post-task questionnaires embedded at the end of each scenario, without interruption. This method works well if your goal is to see participants do all of the scenarios, regardless of the time it takes. It also works well as a strategy if the scenarios are linked, so that the participant finishes one and then naturally moves to the next one.
However, if the participant doesn’t complete one task successfully, you may have to interrupt the flow of the scenarios to help the participant get to the next level. Using this approach means you will need to schedule sufficient time between participants to provide the flexibility to let the session run until all of the scenarios are completed (or decide to stop at whatever point the session needs to end).

**Use these scenario examples for ideas**

The following examples of scenarios build on the earlier examples in this chapter of the characteristics of subgroups. The scenarios are crafted to match the tasks with participants’ goals for the product and your goals for understanding their experience with the product.

**Example: Website for students interested in distance learning**

*Scenario 1: Take a look at the website*
Open up this website (minimized on the tray at the bottom of your screen) and tell us:

- your first impressions
- what you think the site is about
- what you think you might want to do here
- without actually clicking on anything yet, what you think you might want to click on first
- what result you think you would get by following that link

*Scenario 2: Is distance learning right for you?*
You are thinking about the advantages of taking courses online, but you’re not sure if distance learning is right for you.

See if there is information on this site to help you determine whether you are a good fit for distance learning.

*Scenario 3: Finding a degree or course of interest*

*Part A. Finding degree or course*
You have decided you want to find a degree or course that you can take online. According to your questionnaire, you are interested in [field of study]. See if you can find a school that is offering something online that you want to take.

---

*Note:* The scenario does not use “e-learning,” which is the terminology on the website.

*Note:* The task is unstructured to allow participants to pursue a real goal and/or interest. They were screened to match their motivation to the website’s purpose.
Part B. Costs and signing up

Now that you have found something online that suits your needs, do the following:

- Find out how much it costs.
- See if you can sign up for this course or degree.

Example: Digital cable box self-installation

Sometimes, as in the case of the following instructions, the task is a process that fits into a single scenario. In this study, the focus is on the documentation for the self-installation process. The goal of the study is to learn whether the documentation succeeds in providing the support users need to successfully install the digital cable box. An additional goal is to understand when users need to call customer support and, in such cases, whether they are able to complete the self-install after receiving help from someone.

Instructions to participant

Use the documentation to help with the installation, even if you would not normally do this. Remember, we are testing the documentation, not you.

Set up your entertainment system so that the digital cable box, the TV, the DVD player, and the stereo are all connected.

If at any time you feel you need help, call our help desk.

When you are done or ready to stop, give us a call.

Determine quantitative and qualitative feedback methods

Depending on whether you are conducting a formative or summative evaluation and what your goals are for the study, you may want to focus on one type of data collection or another. However, a combination of both quantitative and qualitative collection methods can give you a fuller understanding of the user experience.
If your product is fairly robust, it can be useful and important to set metrics for the study. Management likes metrics because metrics can be used to support the business goals for the product. User experience practitioners like metrics because they can help make a case for product improvements based on usability testing and for more usability testing.

But metrics don’t tell the whole story of your users’ experience. And if your product is at a very early stage of development, metrics may not be appropriate at all. However, when the product is sufficiently developed to take measurements as well as listen and learn from the participants, it’s highly effective to combine metrics with your observations, comments from participants, and their responses to open-ended questions.

For planning purposes, you want to decide which types of data collection you will use. In this chapter, I give you the basics to discuss your options. A few definitions should help shape the discussion.

**Performance and preference data are quantitative**. In other words, the findings can be counted and measured against benchmarks established to determine success or failure.

- **Performance data** are based on measurements of users’ actions, such as time on task; number of errors; recovery from errors; success or failure at task completion; use of help, documentation, or embedded assistance; and so forth. When you set specific performance metrics, you need to base each metric on data. If you arbitrarily set a metric that users must complete the install process within five minutes, their “success” in doing this task in this timeframe may not have any correlation to their perception of ease of use or satisfaction with the process. If you don’t have good metrics before you start testing, you can establish them from your analysis following the first study and then use them as a baseline for future studies.

- **Preference data** are based on users’ responses to questions on post-task and post-test questionnaires. These responses provide quantitative data when they can be measured, using participants’ ratings on tasks (for example, with a five-point scale from very easy to very difficult).

There’s much more about analyzing both quantitative and qualitative data in Chapter 8.
In the planning meeting, you will decide the types of performance and preference data you want to collect. For collecting preference data, you may want to create your own questionnaires, or you may want to use one of the standard questionnaires.

**Observations and user comments provide qualitative feedback**

Observations of your participants yield rich *qualitative* feedback. Qualitative feedback is gathered by noticing what participants do while they are engaged with the product. You may want to note nonverbal feedback, such as users’ facial expressions, body language, and nonverbal utterances, including laughing, sighing, moaning, and even, occasionally, screaming (not to mention cursing!). Don’t underestimate the importance of gathering this feedback: It will tell you a great deal about your users’ experience.

Qualitative feedback also comes from noting participants’ comments as they are thinking out loud while they work and participants’ responses to open-ended questions in a questionnaire or interview after they have completed scenarios and at the end of the testing session. Their candid comments can provide rich insights into their experience, both positive and negative.

**Set dates for testing and deliverables**

Now that you have discussed the goals for your usability study, identified your user subgroup(s), and decided on the tasks to put into scenarios, you have one item left: setting dates for testing and project deliverables.

You may already know the test dates. But it’s often the case that you have a timeframe for conducting testing, but you haven’t set dates yet. Setting the dates for testing may depend on the availability of the core team members, the availability of the lab or space you’ll use for testing, and the discussion you have had in this meeting about the status of the product you want to test.

In addition to setting the dates for testing, you need to decide on the schedule you will use for each testing day. Typical testing takes place...
during normal business hours, but there could be good reasons to test after hours and on weekends, since you need to match the test schedule to the availability of the users.

Another consideration is how to schedule participants from more than one subgroup. If you are doing two days of testing, and if the participants will be easy to recruit, and if you have more than one subgroup, you might want to schedule one subgroup on one day and another subgroup on the other day. This makes the end-of-day analysis a bit easier. But if you don’t have much time and you need as much flexibility as possible, you will probably want to leave participant scheduling up to the recruiter (which could be you!).

Finally, you need to consider the stamina of the moderator in setting your schedule. If you are planning to test with four to six people per day in one-hour sessions, a single moderator should be able to take the schedule in stride. But if you’re planning a longer daily test schedule with more participants, you should consider using more than one moderator. The following sidebar—which is based on feedback from experienced practitioners—provides food for thought on the optimal number of participants a moderator and others on the team can handle per day.

Chapter 6 tells you about recruiting and scheduling participants.

Chapter 7 tells you how to test with two or more moderators.

How many one-hour sessions are optimal for a day?

This is actually a two-part question. The answer to the first part—How many sessions are you good for in a day?—affects the second part—How long does it take to do the analysis? Cliff Anderson, a veteran of usability studies, posed these questions to a professional, private Internet discussion group. His results, based on 30 responses, are shown in the table that follows.
The message to take away from these results is that testing can be quite taxing to the moderator and the team if you try to schedule a lot of participants in a day. But analysis, which comes after testing, is where the time commitment is much greater and the timeframe in which to deliver the results may be much shorter.

<table>
<thead>
<tr>
<th>Question</th>
<th>Mean</th>
<th>Median</th>
<th>Mode</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>How many one-hour sessions are you good for in a day?</td>
<td>4.87</td>
<td>4.75</td>
<td>4</td>
<td>3.5</td>
<td>6.5</td>
</tr>
<tr>
<td>How long does it take to analyze 10 one-hour sessions and write up the results?</td>
<td>39.8 hours</td>
<td>40 hours</td>
<td>40 hours</td>
<td>6 hours</td>
<td>80 hours</td>
</tr>
<tr>
<td>How many years have you been doing user testing?</td>
<td>13</td>
<td>14</td>
<td>2</td>
<td>29</td>
<td></td>
</tr>
</tbody>
</table>

Preparing a test schedule

You may not get the actual schedule done in your planning meeting, but you need to decide the general outline of the schedule so that recruiting can get underway.

What’s in a typical test day? There may be no such thing as a “typical” test day because, as with so much associated with usability testing, “It depends.” If you’re planning to do a walkthrough and/or pilot in advance or on the first day of testing, if you’re planning to do analysis during the day at several break points or at the end of the day, if you’re using the RITE method or another rapid response/redesign approach . . . all of these variables affect your test schedule.

However, a test day, in general terms, has some common characteristics. You’ve always got sessions and (if you’re sane) breaks in between. You’ve always got participants, usually one at a time. You may or may not include time for analysis, prototype changes, or debriefing at the end of the day. But if these activities don’t occur during your test day, they will typically occur as soon as possible afterward.

The ideal test schedule is one that builds in flexibility so that delays can be made up quickly. The reality of a typical test day is that the first
delay can cause a ripple effect throughout the day, made worse by a schedule of back-to-back sessions.

Setting up the schedule that suits your situation

Your schedule needs to accommodate the timeframe for each session, the breaks between sessions, and the time for setup at the beginning and debriefing at the end. If you’re starting with a pilot test, you need to schedule time to make changes afterward. To give you an idea of how to set up your schedule, take a look at this table, which shows you several schedule options for a one-day study of one-hour sessions.

<table>
<thead>
<tr>
<th>Time</th>
<th>Option 1: Debrief at end of day</th>
<th>Option 2: Debrief during the day</th>
<th>Option 3: Pilot plus testing and debrief</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:00–9:00 A.M.</td>
<td>Setup/preparation</td>
<td>Setup/preparation</td>
<td>Setup/preparation</td>
</tr>
<tr>
<td>9:00–10:00 A.M.</td>
<td>Participant 1</td>
<td>Participant 1</td>
<td>Pilot test</td>
</tr>
<tr>
<td>10:15–11:15 A.M.</td>
<td>Participant 2</td>
<td>Participant 2</td>
<td>Pilot debrief/revision</td>
</tr>
<tr>
<td>11:30–12:30 P.M.</td>
<td>Participant 3</td>
<td>Debrief</td>
<td>Participant 1</td>
</tr>
<tr>
<td>12:30–1:30 P.M.</td>
<td>Lunch brought in</td>
<td>Lunch brought in</td>
<td>Lunch brought in</td>
</tr>
<tr>
<td>1:30–2:30 P.M.</td>
<td>Participant 4</td>
<td>Participant 3</td>
<td>Participant 2</td>
</tr>
<tr>
<td>2:45–3:45 P.M.</td>
<td>Participant 5</td>
<td>Participant 4</td>
<td>Participant 3</td>
</tr>
<tr>
<td>4:00–5:00 P.M.</td>
<td>Debrief</td>
<td>Participant 5</td>
<td>Participant 4</td>
</tr>
<tr>
<td>5:00–6:00 P.M.</td>
<td>Debrief</td>
<td>Debrief</td>
<td>Participant 5*</td>
</tr>
<tr>
<td>6:00–7:00 P.M. or later</td>
<td>Wrap up or prepare for next day</td>
<td>Wrap up or prepare for next day</td>
<td>Debrief or wrap up/prepare for next day</td>
</tr>
</tbody>
</table>

*Start time is 5:15; debrief begins at 6:30.

This schedule shows 15 minutes between participant sessions. You could expand it to give yourself more time between sessions, but it’s not a good idea to compress it, since you really need the flexibility of at least 15 minutes between sessions to allow for any resetting you need to do, such as reassembling the paper prototype, clearing the cache on a website or application, or deleting the account you had the user create.

Another variation to this schedule is that you could start earlier or later and go into the evening, particularly if you need to schedule evening sessions when your participants are available.
You could also set the debrief for another day, so that you can maximize your testing time with participants and see one or two more per day. If you are paying for a lab rental and you want your observers to see as many participants as possible, this could be an important consideration.

If you conduct the debriefing/analysis session at the end of the day, it can be more than an hour or two, although it’s good to set a time limit on this activity (with the option to continue another day or in different medium).

**Setting dates for deliverables**

The next item related to setting dates is determining when the draft deliverables will be circulated for review and when the final deliverables will be completed. In addition, you need to decide who will take the lead on each of these deliverables. As I said earlier, if testing is going to be soon, the first deliverable needs to be the screener, since screening for the right participants has the longest lead time.

The due date for the deliverables will be determined by your plan to test the test. If the product and the testing facility and the team are available before full testing begins, you will need the deliverables for a walkthrough or pilot. If it’s not feasible to test the test in advance, then the deadline for the deliverables may be the day before testing starts.

Finally, you want to agree on the post-test deliverables so that everyone understands when results will be available and in what form.

**Writing the test plan**

No matter how much buy-in you get for everything you have planned for your usability study, you still need to document it. Whether informal or formal, the test plan is the record of the decisions made about what to test, how to test, who to recruit for testing, and so forth. Without this document, individual memories may vary, decisions may become blurred, and the outcome from testing could be challenged.

The test plan puts everything in writing. If the planning meeting gave you the time you needed to produce drafts for all of the materials to be used in the test, then the test plan can be written immediately after the meeting. More common, however, is to continue working on materials for the test after the meeting. In this situation, you should think of the test plan as a living document that evolves as the materials get fleshed out.
It’s best to determine a freeze date for the test plan, though, so that everyone can agree to work with the same plan, based on the last revisions allowed.

The test plan can be informal or formal, depending on the needs of the recipients of the plan. The type of test plan people expect should be decided in the planning meeting.

Writing an informal test plan

An informal test plan can be nothing more than the notes or minutes from the planning meeting. The three pages of Figure 5.4 show the notes from a meeting to plan the usability study of FTP software. I have used parts of the planning process for this study as an example in this chapter.

---

**Planning Meeting Notes**

New version, release xxx
Two user groups—novice and advanced
Meeting participants: (list)

**Date of meeting:** January 6, 20xx
**Meeting time:** 1:00–5:00 P.M.

**Product goals:**
- increased conversion rate following product release
- reduced support calls

**Test goals:**
- confirm whether the install is simplified for new users
- learn what new users do when error occurs
- learn new users’ expectations from name, download, install experience
- learn advanced users’ perceptions of new product (compared to other products)
- learn advanced users’ perceptions of new features (e.g., PGP)
- understand the experience for the new user in the purchasing model of try/buy

**Status of product for test/issues:**
New version of current product: mature product; not done formal testing before; many features added to product over time; lesser-skilled users are now using the product. Testing will take place at two points in development so that changes resulting from first test can be tested in a follow-up study.

**Issues in current product, to be studied and improved in new release:**
- initial screen is confusing
- product doesn’t have modern look and feel
- confusion over two interface options in the product; which to choose?
- problems with file transfer manager opening in new window

---

**Figure 5.4** These notes from the planning meeting document the decisions made for testing the FTP software.
New product features include:
- multiple interfaces (MDI type model); goal is to be more flexible
- file transfer manager embedded in the interface (not separate window)
- log-in embedded in interface
- Internet Explorer embedded as a choice for interface design

User profiles [these were developed from the characteristics of the user subgroups]:
Two user groups: novice and advanced (screening questionnaires to be drafted by assigned team member)

Novice user:
- must have home computer with network connection
- minimum one year Internet experience
  - minimum ½ hour per day Internet use (or 4 hours per week)
  - surf Internet for information, shopping
  - use e-mail, attachments
- must have downloaded information from Internet
- technical abilities can include use of scanner, digital camera, MP3, or downloaded music
- no network concept understanding (terms such as host and IP address are likely to have no meaning)
- must not have had experience building a website (except, perhaps, using a wizard to create a web space at a portal site like Yahoo!)
- must have interest in creating a website or using a host site to share or post files (example, eBay)
- must not have had experience loading files to an FTP server
- age not an issue
- must be native/fluent English speaker (no ESL speakers)

Advanced user:
- current FTP user, but not a current or recent customer (may select users whose experience with product is at least two versions earlier)
- likely to be Webmaster, IT manager, or IT professional
- looking to gain productivity
- has had FTP experience with one or more other FTP products: (list to follow)

Questions to ask potential participants:
- How long and how often have you used FTP products (for each of the products on the list)?
- How would you make a second simultaneous connection?
- Do you manage encrypted files? If so, how?

Tasks for Scenarios (to be drafted by assigned team member):
Novice user:
1. Initial impressions—install (five screens). After installation wizard is completed, ask user for impressions of process, product, name of product, etc. Also find out perceptions of options and functions.
2. eBay or similar site. Connect to the site. Upload a file. Add another file.
3. Download file(s) from public site like NASA (Mars photos).

Figure 5.4 These notes from the planning meeting document the decisions made for testing the FTP software. (Continued)
Advanced user:
1. Launch product. Upload all files to one server.
2. Create connection to second server. Add sites.
3. From local server, upload files to two other servers. If they do this in version A, ask them to find another way to do it (to discover the differences with Version B).
4. Your company has just made a policy that all files shared on servers must be encrypted using PGP. Encrypt your local files; provide access for four people.

Performance measurements:
- help (use and types)
- errors and recovery
- time on tasks
- satisfaction
- learnability
- ease of use

For advanced users: Comparison with other FTP products

For both user groups:
- Would you purchase the product?
- How much would you be willing to pay for it?

Dates for deliverables and testing:
- screening questionnaires completed by Jan. 9; recruitment begins
- product prototype to team members for scenarios; freeze date on prototype for the test Jan. 13
- all materials for walkthrough must be reviewed by team and completed by Jan. 13
- walkthrough (two versions of scenarios for two user groups), Jan. 15, 11:00 A.M. to noon; 1:00 P.M. to 2:00 P.M. (lunch provided; revisions made over lunch and after second user, as needed)
- revision to materials as needed; finalized by Jan. 16
- pilot Jan. 19, 1:00 P.M. to 3:30 P.M. (two users—novice and advanced—back to back)
- revision to materials as needed, same day
- testing dates (specific schedule for recruiting, sent separately)
  - day 1, Jan. 20, 8:00 A.M.–4:30 P.M. (setup, five participants)
  - day 2, Jan. 21, 8:00 A.M.–7:00 P.M. (setup, four participants + replacement, if needed; debrief)

Figure 5.4 (Continued)

Writing a formal test plan
A formal test plan is produced when full reporting is required or expected. Even when not required, a formal test plan can be helpful when key stakeholders are not present at the planning meeting.

A formal test plan typically contains the following sections:
- **Title Page**—identifies the document as the test plan for a particular product/group/company/date(s). The plan is directed to the sponsor
(or key decision maker) for the study and identifies who the plan is from, either the team lead or the group or external consultancy.

- **Table of Contents**—reflects the first- and second-level headings in the test plan, with corresponding page numbers.

- **Purpose/Executive Summary**—a purpose statement provides an overview of the test plan and the purpose of the test. An executive summary is more explicit, providing management with a succinct but clear description of the critical elements of the test plan. This is usually a one-page summary that executives and managers can read quickly and know the essential elements of the plan.

- **Problem Statement and Test Objectives**—establishes the issues to be addressed in the test, framed as goals.

- **Methodology**—describes the type of test and the method to be used.

- **User Profiles**—describes the users for this test. If the test will address two or more subgroups of users, then each profile provides the specific characteristics.

- **Participant Incentive**—defines the amount or type of incentive to be used in recruiting participants.

- **Screeners**—created for each user profile. Can be included in the report or in the appendix.

- **Task List**—describes the tasks that will be included in the test, sometimes with the objective of each set of tasks included to show how these tasks match the study's goals.

- **Scenarios**—presents the tasks within realistic, goal-directed descriptions. The tasks and goals can be included at the top of each scenario. Can be included in the report or in the appendix.

- **Evaluation Methods**—describes the data collection methods, including the types of data that will be collected (quantitative and qualitative). If questionnaires have been developed or identified at this point, they are included.

- **Test Environment and Equipment**—describes the equipment in the facility if it is not known to the report's readers. Also, if the testing environment needs to be configured in a certain way or requires additional equipment, that information is documented here.

- **Deliverables**—describes the reports that will be delivered following testing. Also describes the method and type of delivery—formal or informal, paper or electronic document; oral presentation/meeting; video highlights tape—and dates for delivery.
Summarizing Chapter 5

This chapter presented the steps you need to take to plan your usability test. Those steps become the agenda for the planning meeting you will have with the key stakeholders to decide on the elements of your usability test. The agenda items for the meeting include:

- setting test goals by deciding what’s most important to learn from users in this study
- determining how to test the product by discussing what to test, where to test, and how to set up the test
- agreeing on the user subgroup(s) you will use by defining the characteristics of each subgroup and putting these characteristics together in user profiles
- determining the appropriate incentive for each subgroup
- drafting the screeners based on the characteristics of the user subgroups
- selecting the tasks to put into scenarios by matching your goals with your users’ goals and crafting scenarios that feel “real” to your users
- determining the quantitative and qualitative feedback methods you’ll use to capture metrics and observations
- setting dates for testing and deliverables for the test and the test results

With all of these decisions made in the planning meeting, and with deliverable dates set for completing the elements for the usability test, you now want to document your decisions. That’s what the test plan does.
Test plan for Holiday Inn China website usability study

The test plan for the Holiday Inn China website usability study was presented in a formal report because it was going to the sponsor/client, the User Experience manager at Intercontinental Hotels Group, for review by the manager and others. The plan contains several appendixes, which include questionnaires to be used in testing sessions. Because every part of the test had to be approved by the sponsor, these parts are included in the test plan.

Part of the test plan is included here. To see the complete test plan, visit the book's companion website at www.mkp.com/testingessentials.
Usability Test Plan for HolidayInn.com.cn

Holiday Inn Hotel Website for Chinese-Speaking Users

Prepared for: Karen Bennett, Project Sponsor, User Experience Manager at IHG

Cc: Dr. Carol Barnum, Project Advisor

Completed by: Team CBR
Yufei Duan
Yina Li
Ying Li
Qianying Liu
Niven Sellars
Michael Somer

Date: March 02, 2008
Purpose

The purpose of the *holidayinn.com.cn* usability test is to collect feedback about how users use the Chinese Holiday Inn website, what problems they may encounter using the site, and what improvements they would like to see to make it easier to book a hotel room.

Ms. Karen Bennett, the User Experience Manager of InterContinental Hotels Group, is the project sponsor. She would like to learn about users’ experiences on the Chinese website compared to their experiences with a competitor website: *www.elong.com*. If time permits, she would also like to know how their experiences compare with *www.holiday.com*, the U.S. website.

This test plan describes:

- problem statement and test objectives
- user profile to recruit for testing
- testing methodology and tasks/scenarios
- test setup in the usability center
- plans for data collection and reporting
- project deliverables
- questionnaires and other materials to be used in testing, in appendices:
  - Appendix A: Participant screening questionnaire
  - Appendix B: Pre-test questionnaire
  - Appendix C: Post-task questionnaires
  - Appendix D: Post-test questionnaire

Problem Statement and Test Objectives

This usability test of the Chinese Holiday Inn website will provide qualitative and quantitative data addressing IHG’s interest in understanding Chinese users’ experience. We will assess the users’ experience with booking a room as well as their satisfaction with the site. Tasks for testing *www.holidayinn.com.cn* include:

- *The general feeling/layout of the site:* Does the layout suggest the route first-time users will take to book a hotel?
The procedure for booking a hotel online:

- **Basic search:** Is it easy to use?
- **Advanced search:** Can users accomplish their goals on the advanced search screen?
- **Entering personal information:** Does the website require reasonable and suitable information for Chinese users? Do users understand all information requirements?

**Language:** Do users understand all the wording on the website? Are there any translation mistakes or misunderstandings?

**Information in confirmation e-mail:** How quickly can users receive the confirmation e-mail after booking a hotel online? Does the confirmation e-mail contain enough information to suit users’ needs?

**Perceived reliability of the site:** Do users trust the website? Do they fill out the personal information readily?

**Navigation:** Can users find the most efficient navigation when they book a hotel or browse the website?

**Satisfaction:** Which aspects do users like and which aspects do they dislike?

These questions were devised using information supplied by our sponsor and our assessment of potential usability problems resulting from our heuristic evaluation of the site (submitted previously).

This usability study will be designed to allow us to obtain mostly qualitative data. The focus of the study will be task-oriented and directed toward how the user subjectively responds to the issues listed above. Participants will be given a pre-test questionnaire to get information about their hotel booking experience and expectations, followed by scenarios that direct them to perform specific tasks. Post-task questionnaires after each scenario, as well as a post-test questionnaire, are designed to obtain detailed and specific feedback about [www.holidayinn.com.cn](http://www.holidayinn.com.cn) features.

Users will be scheduled in one-hour sessions to include pre-test and post-test questionnaires. Users will be given 40 minutes to complete five scenarios. If participants complete the scenarios in less than 40 minutes, we have designed an extra scenario to obtain additional feedback.
The test will be conducted in a full-scale usability lab at Southern Polytechnic. Logging and recording of sessions will be done using Morae, with backup recordings on DVD. The participant computer will have Microsoft’s Chinese character-based language installed.

**User Profile for U.S.-Based Participants**

To recruit prospective test participants who represent the site’s actual users, the project team identified two primary user groups and created two descriptions of those users. Personas for these two user profiles were submitted previously.

Since the situation is that of testing the Chinese website within the United States, we will use only one of the user groups/personas, which will allow us to recruit these users to participate in the study in the Usability Center.

The characteristics of the U.S.-based user group, along with other questions about prior experience, behaviors, and other criteria, will make up the screening questionnaire (Appendix A) for identifying suitable test participants.

A general description of the user characteristics is as follows:

- must be able to speak and read Chinese as well as English
- first language must be Chinese
- travel purpose is either business or pleasure
- age must be between 24 and 55
- need a mix of female and male
- must have experience booking a hotel room online
- must use the Internet at least five hours per week
- must not have any prior experience with the Chinese Holiday Inn website

**Methodology and Tasks/Scenarios**

This section describes the testing methodology and scenarios we will use with six participants in one-hour sessions.
Number of Participants
We plan to test a total of six participants: three who travel for business and three for pleasure. To ensure an adequate number of participants, we plan to recruit eight users based on our screener; two out of the eight are backup participants.

Length of Sessions
The total length of each session will be an hour; including:

- Welcome and pre-test questionnaire: 10 minutes
- Task scenarios: 40 minutes
- Post-test questionnaire: 10 minutes

The estimated time for each scenario is noted below.

Test Procedure
The test will begin with an overview briefing, followed by the scenarios described below, post-task questionnaires, and a post-test questionnaire.

Participants will not be allowed to use any resources that are not available on the Holiday Inn website (that is, no use of Google or other outside sources). This restriction will not be stated in the moderator’s introduction to avoid influencing the participants’ actions. However, if participants want to use external resources, the moderator will notify them that such action is not allowed for testing purposes.

Overview/Briefing (10 minutes)
The moderator will welcome the participant and have him/her sign the video consent form. Next, the moderator will explain the facilities and ask the participant to think out loud. Also, the moderator will provide the participant with the pre-test questionnaire (Appendix B).

Scenario 1 (5 minutes)
You and your spouse are thinking of going back to China for the Summer Olympics. You heard that the Holiday Inn has several locations in Beijing.
You will visit the Chinese Holiday Inn website. (This will be minimized on the desktop.) Take a moment to look at the homepage without clicking on anything. After you have familiarized yourself with the homepage, tell us: What do you think you would do on the site to book a hotel room?

**Scenario 2 (15 minutes)**
You are going to book a hotel room for your stay in China. You want to reserve one hotel room with one bed for two people in Beijing. You will stay from July 20 to July 25. Here is the information you need to book the room:

- **First name**: Jing
- **Last name**: Li
- Reserve for you and your spouse
- **Location**: Beijing
- **Check-in date**: July 20, 2008
- **Check-out date**: July 25, 2008
- **Credit card type**: Visa
- **Credit card number**: 1234 2345 3456 4567
- **Credit card security code**: 990
- **Credit card expiration date**: 10/2010
- **E-mail**: holiday@gmail.com

**Scenario 3 (5 minutes)**
You just answered a phone call from your relative who lives in Beijing. She wants you and your spouse to stay at her house during the visit. You accept the invitation, so you must now cancel your Holiday Inn reservation.

**Scenario 4 (5 minutes)**
You have some relatives in Changchun, Jilin Province. You want to visit them during your trip to China. See if there is a Holiday Inn available in Changchun from July 26–30.
Scenario 5 (10 minutes)

Go to the competitor site www.elong.com. See if a hotel room is available for you and your spouse between July 21 and 25 in Beijing. (Note: You are not going to actually book a hotel room on this website.) Look at the search results page and tell us:

- What do you think of this site?
- What do you like about it?
- What do you dislike about it?

Scenario 6 (optional; 5 minutes)

- Go to the Holiday Inn U.S. website (www.holiday.inn.com).
- Take a moment to look at the homepage without clicking on anything.
- Tell us what you think of this website compared to the Holiday Inn China and eLong websites.

Closing (5–10 minutes)

Moderator will ask the participant to complete the post-test questionnaire (Appendix D).