5

Running a Research Session

This chapter provides practical information on what to do when you run a research session. We assume that you have developed your initial experimental design and are now ready to run a pilot study. This chapter is thus about interacting with subjects and the context in which you do that.

There are two main sections, preparing for a session and running the session. Accompanying figures summarize the steps. This chapter also notes other issues that can arise while running a study, including computer simulations (models) as subjects, missing subjects, and other problems.

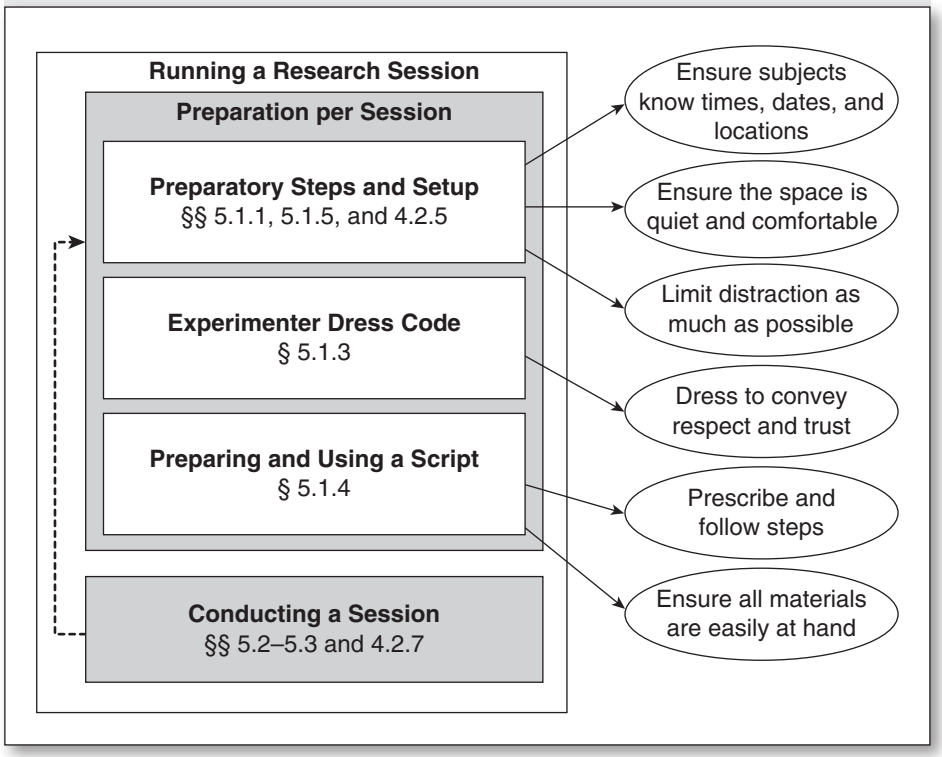
5.1 Preparing to Run a Research Session

There are several steps for preparing a study. Some of these steps are done much earlier, such as piloting the study to prepare in general, and some are done on the day of a session to prepare for the next subject. Figure 5.1 provides a graphical overview of the process of preparing a study.

5.1.1 Preparing the Space for Your Study

The environment you provide for your subjects directly influences the quality of your data. Typically, preparing the space for your experiment will seem straightforward—often, subjects will simply sit at a computer and perform the experimental task. However, giving some thought to setting up the space in advance can help. For example, if possible, you should provide an adjustable-height chair if subjects are sitting at a computer. Minimizing

Figure 5.1. A pictorial summary of preparing a research session, along with the section (§) or sections (§§) that explain each step.



screen glare from overhead lights can be important—it may be helpful to use an incandescent table lamp instead of bright fluorescent ceiling fixtures. Allow for the possibility that some of your subjects may be left-handed—we have seen experimental setups that were very awkward for left-handers to use. In general, try to take the perspective of your subjects and make the setup as comfortable as possible for them.

In setting up the space, it is also important to consider possible distractions. For example, if your experimental space is next to an office or opens on a busy hallway, consider the possibility that loud conversations nearby may distract your subjects. The ideal setup for running individual subjects is a sound-isolated chamber or room, but such a space is not always available. A simple sign that reads "Experiment in Progress—Quiet Please" can help a great deal. If you must collect data in a room used for other purposes, such a sign can also help avoid accidental intrusions by others who may not realize an experiment is in progress. (Also, take the sign down after the study,

or people in the building will learn to ignore it.) It is also best to remove "attractive nuisances"—objects that invite inspection—from the experimental space. For example, one of us collected data in a room that had a shelf full of toys and puzzles used in another study—until we found a subject playing with a puzzle rather than performing the experimental task!

Often, subjects may have to wait after arriving at your study, perhaps as other subjects finish. While you should try to minimize waiting time—unlike a doctor's office or driver's license center, your subjects don't *have* to be there—it is important to provide a comfortable place to wait. If the only waiting area available is a hallway, try at least to place chairs in an appropriate location with a sign that says "Please wait here for [TitleOfExperiment] experiment."

Figures 5.2 and 5.3 show two spaces used for running subjects in a psychology department. Figure 5.2 shows a small storage space used as a single-subject data-collection station. A table lamp is used to avoid glare from overhead fluorescent lights, and the room is kept free of distractions. The room is on a quiet, rarely used hallway, so this space provides good isolation. A nearby workroom serves as a reception and waiting area, as well as office space for research assistants. If there is not an easy way to watch the subjects through a one-way mirror, we have found video baby monitors to be helpful.

Figure 5.2. A storage space used as a single-subject data-collection station.

Figure 5.3. An office space used to house multiple data-collection stations.

Figure 5.3 shows a large office used to house multiple data-collection stations. Office dividers separate the stations and provide some visual isolation while allowing a single experimenter to instruct and monitor several subjects simultaneously. In such setups, subjects are sometimes asked to wear headphones playing white noise to provide additional isolation. In this space, subjects wait for their sessions in the hallway, which also has a sign asking for quiet.

5.1.2 Piloting

As mentioned earlier, conducting a pilot study based on the script of the research study is important. Piloting can help you determine whether your experimental design "works." If a revision to the study is necessary, it is far better to find it and correct it before running multiple subjects, particularly when access to subjects is limited. It is helpful and far less stressful to think of designing experiments as an iterative process characterized by a cycle of design, testing, and redesign, as noted in Figure 1.1. In addition, you are likely to find that this process of running an experiment works in parallel with other experiments and may be informed by them (i.e., lessons learned from ongoing related lab work may influence your thinking).

Thus, we highly recommend that you use pilot studies to test your written protocols (e.g., the instructions for experimenters). The pilot phase provides experimenters the opportunity to test the written protocols with practice participants and is important for ironing out misunderstandings, discovering problematic features of the testing equipment, and identifying other conditions that might influence the participants. Revisions are a normal part of the process; do not hesitate to revise your protocols in consultation with the principal investigator (PI). This will save time later. There is also an art to knowing when not to change the protocol. Your PI can help judge this!

The major reason for returning to the topic of piloting is that the pilot study provides an opportunity to think through the issues raised here—the setup of the experimental space; interacting with subjects before, during, and at the conclusion of the experiment; and so on. Especially for an inexperienced experimenter, pilot testing provides an opportunity to practice all these things. In some cases, it may be effective to begin pilot testing with role-playing—one member of the research team plays the role of the subject, while another plays the role of experimenter.

You will often start piloting with other experimenters and then move to officemates and people down the hall. One researcher we know gets Institutional Review Board (IRB) approval early and switches to subjects that can be kept, using them as pilot subjects. When the process is smooth, this researcher declares them keepers. This is expensive, but for complicated studies is probably necessary because your lab mates know too much to be useful pilot subjects. It is important to keep in mind that once you involve actual subjects whose data you may keep, or who are recruited from a subject pool, all the issues concerning IRB approval discussed earlier come into play.

It is also important when piloting to test your data-gathering and analyses steps. We have wasted significant amounts of resources when the apparatus did not measure what we thought it did, and we know of numerous studies where the format of the study software output did not load easily and directly into analysis software or did not record the information that was later found to be needed. So, as an important part of piloting, take some of the pilot data and test-analyze them to see that the data are recorded cleanly and correctly, that they load into later analysis tools, and that the results you want to examine can be found in the recordings you have. You can also start to see if your manipulations are leading to changes in behavior.

5.1.3 Experimenter Dress Code

The goal of a dress code is to convey a serious atmosphere and to encourage respect and cooperation from your subjects. You should consider the impression you wish to make and will make when running your experiment.

This consideration should include how you wish to position yourself (in a way that commands respect while making the participants comfortable enough to perform the task), the type of experiment, and the type of participants in the experiment.

In most cases, we recommend wearing a semiprofessional outfit ("office casual"), such as a dress shirt with dress slacks, when running experiments. This helps you look professional and prepared but not intimidating. Semiprofessional dress helps convey the experiment's importance without overwhelming the participant. However, appropriate dress may vary depending on your subject population. If you are a college student interacting with college-student subjects, it may be best to dress like a college student—but think of a college student who wants to make a good impression on a professor, not a college student hanging out in the evening. It is certainly best to avoid things like T-shirts with slogans some might find offensive, low-cut blouses, very short shorts or skirts, or flip-flops. If you are working with non-student adult subjects, business casual is a better choice of dress. If your subjects are expert professionals, you should dress in a way that would fit their workplace.

5.1.4 Preparing and Using a Script

Your research study will likely have a script of how to run the session. If it does not, it should, because a script will help you run each subject in a confident and consistent manner. The script will often start with how to set up the apparatus. Before the subject's arrival, the experimenter needs to set up the apparatus and should be ready to welcome the subject. Incorrectly or inconsistently applied procedures cause inconsistencies in running the experiment (e.g., omission of an instruction resulting in noisier data). Consequently, the script that appropriately represents required procedures plays an important role in conducting a successful experimental study. Appendix 2 provides an example study script.

The setup should include making sure all materials that will be used in the session are available (e.g., forms, at least one backup copy) and that the apparatus is working. If batteries are used for any part of the apparatus (e.g., a laser pointer, a DVD remote), spare batteries should be on hand.

5.1.5 Before Subjects Arrive

Your interaction with the subjects you've recruited begins before they arrive. It is wise to remind subjects by phone or e-mail the day before a study

is scheduled, if they have been scheduled farther in advance, and to repeat the time, place, and directions in the reminder. If there is a time window beyond which you cannot begin the study—for example, you might need to exclude from a group study anyone who arrives more than 5 minutes late—make sure this is clear as well.

As you schedule the times to run, you should take advice about when to schedule times. It is usually appropriate to schedule times during normal business hours (which in a university lab may be 10 a.m. to 6 p.m.). If you are running subjects outside of these normal hours you should have a discussion with the PI about safety for you and the subjects (how to reach the PI, for example). You should also consider practical issues such as whether the building will be locked after normal business hours or on weekends. If your subjects are traveling some distance to be in your experiments, do parking restrictions or bus schedules change after hours or on weekends?

It is also important to be clear about where the study is. Make sure that your subjects have clear directions to the location of your study. On a college campus, it may be important to provide directions and identify nearby landmarks. If subjects are driving to the location of your study, make sure you provide clear instructions on where to park and whether they are expected to pay for parking. Make sure the door to the building is unlocked, or have someone meet subjects at the door—one of us knows of an experiment in which several subjects were not run and hours of data collection were lost because the experimenter didn't realize the campus building would be locked after 5 p.m. and the subjects were literally lost.

You should also provide clear directions to the specific room in which the study is held. Several of us have seen research subjects wandering the halls looking for the room their experiment is in. It is also helpful to clearly mark the place where the experiment will be (or the place where subjects should wait)—a simple sign that says "Skill Acquisition Experiment here," for example, may save a lot of confusion in a building where all the halls and doorways look pretty much alike and where multiple experiments are in progress, or in buildings where the rooms are not numbered appropriately. If subjects must pass a receptionist to find your study, make sure the receptionist knows where the study is and who is running it—many people will stop to ask even if they think they know where they're going.

Making it as easy as possible for subjects to find your study and to arrive in a timely manner is important to ensure that they arrive ready to participate, with minimal anxiety. This helps in establishing the cooperative relationship with your subjects that will yield the best results for your experiment.

5.2 Running a Research Session

A research session will have roughly four phases: welcoming the subject, running the study itself, debriefing the subject, and then providing the subject with compensation when promised. Figure 5.4 provides a graphical overview of running a research session.

5.2.1 Welcome

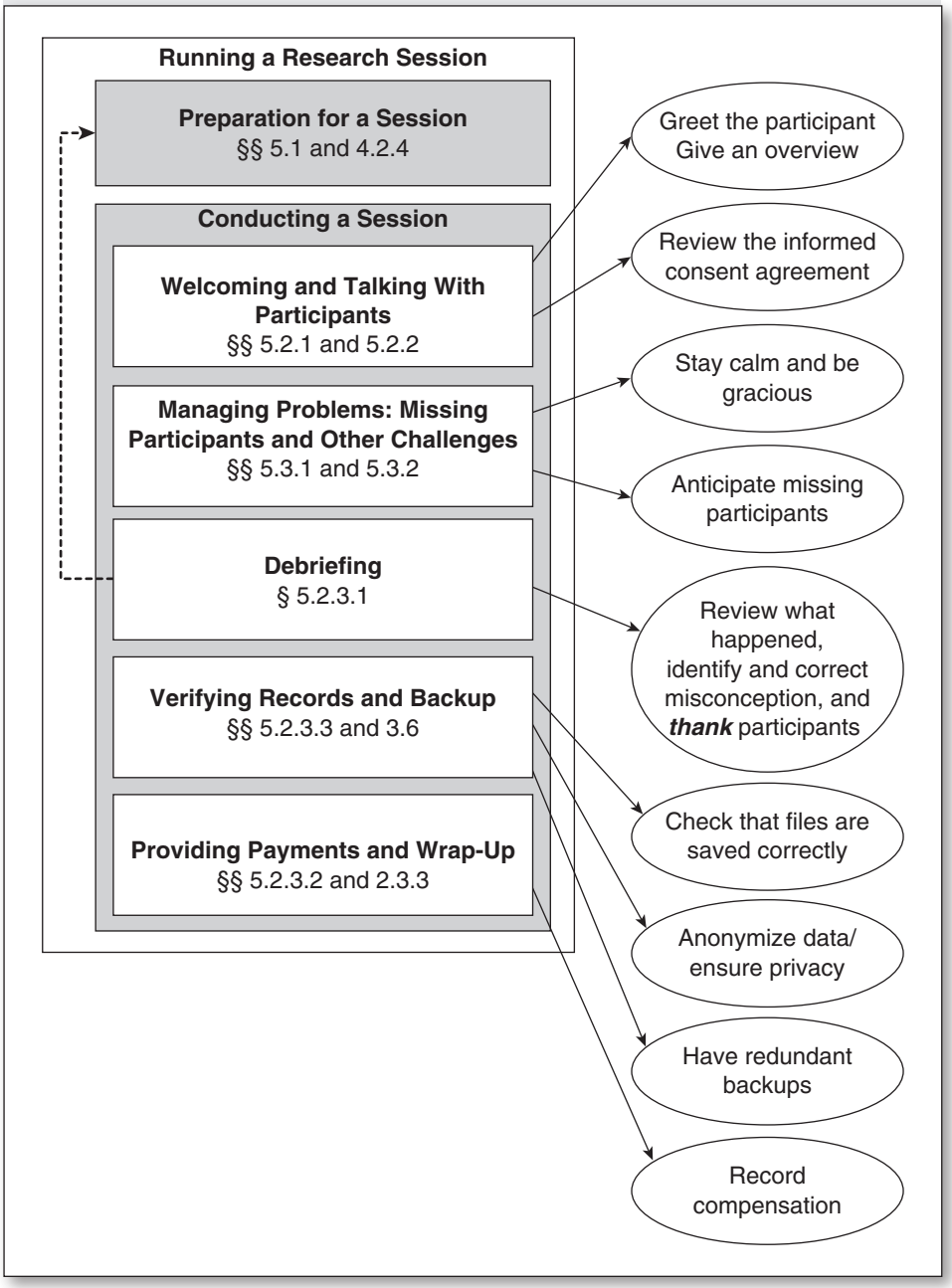
As the experimenter, you are taking on a role similar to that of a host; thus, it is appropriate to welcome participants to the study. You might provide them reading materials if they have to wait and should answer any questions they have before the study begins. It is also appropriate to confirm their names (for class credit) and to confirm for them that they are in the right place at the right time. If the experimental protocol permits it, you might also indicate how long the study will take. This helps set the stage for the study itself.

The first event after welcoming subjects is typically the informed consent procedure. It is important to take this seriously—while it will become routine to you, it is likely not routine for your subjects. Rather than simply handing a subject the consent document and saying, "You have to sign this before we can start," take the time to explain the major points and to provide an opportunity for questions. Many will have no questions and will glance quickly at the document before signing it. Nevertheless, your approach every time should be one that allows the subject an opportunity to understand and think about what he or she is agreeing to.

5.2.2 Talking With Subjects

When you welcome the subjects to your study and the study area, you might feel uncomfortable in the first few sessions you run. After you have run a few sessions, this discomfort will go away. In a simple study, you can be quite natural, as there is nothing to "give away." In more complex studies, you will be busy setting up the apparatus, and this tends to make things less stressful for you. It is important, however, to realize that talking with subjects before they begin the experiment plays an important role in getting good data. Often, subjects come to the lab feeling nervous, with little or no experience participating in research and, perhaps, with misconceptions about the nature of behavioral research. For example, it is not unusual for students participating in university subject pools to believe that all experiments involve deception or that all researchers are surreptitiously evaluating their

Figure 5.4. A pictorial summary of running a research session, along with the section (§) or sections (§§) that explain each step.



personalities or possible mental disorders. Interacting in a natural, cordial way and explaining clearly what your subjects will be asked to do can go a long way toward alleviating the subjects' anxiety and ensuring that they do their best to comply with the instructions and complete the experimental task. In our experience, it is all too easy for experimenters to interact with subjects in a rote and too remote manner that increases rather than alleviates their anxiety. Remember that although you may have repeated the experimental protocol dozens of times, it is the first time for each subject!

In nearly all cases, abstaining from extraneous comment on the study is an important and useful practice that makes all parties concerned more comfortable. Many experimental protocols require not giving the subject feedback during the study. In these cases, your notes will probably indicate that you tell the participants at the beginning of the session that you are not allowed to provide them feedback on their performance. Generally, the debriefing can handle most questions, but if you are not sure how to answer a question, either find and ask the PI or take the subject's contact information and tell him or her you will get an answer. And then do it! This also means that when you are running subjects for the first couple of times, someone who can answer your questions should be available.

In social psychology studies or where deception is involved, you will be briefed by the investigator and will practice beforehand. In this area, practice and taking advice from the lead researcher is particularly important.

Be culturally sensitive and respectful to the participants. Consult with the lead investigator if you have general questions concerning lab etiquette or specific questions related to the study.

There are a few things that seem too obvious to mention, but experience tells us that we should bring them up. Don't ask a subject for his or her phone number, no matter how attractive you find the subject! The experiment is not an appropriate context to try to initiate a romantic relationship. Don't complain about how hard it is to work in the lab or how difficult you found your previous subject. Don't tell a subject that his or her session is the last session of your workday, so you hope the session is over quickly. And so on. It might seem that nobody with common sense would do any of these things, but we've seen them all happen.

5.2.3 Concluding a Session

After your subject has finished participating in your experiment, several important parts of your interaction with him or her remain to be completed. These include debriefing, providing compensation, and checking the data.

It is also wise when concluding the experiment to make sure you have all the information you need from the subject. Do you have your copy of the consent document signed by the subject? Is information that will allow you to link pencil-and-paper data with computer data files properly recorded?

5.2.3.1 Debriefing

The American Psychological Association's ethical principles offer a general outline of debriefing procedures. For many experiments, the lead researcher may provide additional guidance. Experimenters should ensure that participants acquire appropriate information about the experiment—such as the nature, results, and conclusions of the research. If participants are misinformed on any of these points, investigators must take time to correct these misunderstandings. Also, if any procedures in a session are found to harm a participant, the research team must take reasonable steps to report and alleviate that harm.

Reviewing your plans for debriefing will be part of obtaining approval for your experiment from the IRB or ethics panel. Sometimes, there are local rules about debriefing—for example, a university subject pool may require a deeper educational debriefing for every study, even when the IRB does not. In an educational debriefing, you would describe the design of the study and the theoretical question it addresses in more detail, using keywords that allow the subjects to see connections between participating in your study and what they are learning in their classes. You may be required to provide a written debriefing or to have your debriefing approved by the administrator of your subject pool.

The experiment's procedures may cause participants to feel uncomfortable or alarmed (although this would be unusual). After the experiment is finished, investigators or experimenters should listen to the participants' concerns and try to address these problems. Mitchell and Jolley (2012) provided reasonable steps to follow when you debrief:

a. Correct any misconceptions participants might have.
b. Give a summary of the study without using technical terms and jargon.
c. Provide participants an opportunity to ask any questions they might have.
d. Express gratitude to the participants.

As with the informed consent procedure, you may find that some, even most, subjects are uninterested in the debriefing. Also, debriefing will become routine to you as you run more subjects. It is important not to let these things lead you to approach debriefing in a perfunctory way that conveys to all subjects that you do not consider it important. If only one subject appears interested, that is reason enough to take debriefing seriously.

As noted earlier, if your study has involved deception, you must usually reveal this deception to the subject. Even if there was no deception, it is good practice to spend a few minutes debriefing the subject about the purpose of the study—your hypotheses, how you hope to use the results, and so on. These are things you generally don't want to mention at the beginning of the experimental session but that will help your subjects understand the value of their participation.

When you have a study that can be perceived as being deceptive or when the study is a double-blind study, you should seek advice about how to debrief the participants. If deception is a procedural component, you will most likely have to explain this to the subject and ask that the subject not discuss the study until all the subjects have been run (after the study's completion date). For all studies, requesting that the subject refrain from discussing the study will help keep potential subjects from becoming too informed.

To review, double-blind studies prescribe that neither the subject nor the experimenter knows which treatment the subject has received. For example, the amount of caffeine a subject has ingested in a caffeine study with multiple possible doses would be revealed to the subject by a third party after the subject had left the experimenter, and this information would not be examined by the experimenters until all the subjects had been run. In these cases, you will have to explain the procedures of the study as well as provide a general rationale for double-blind trials. Otherwise, participants may balk at being given a treatment in a sealed envelope or by a person who is not the experimenter. Furthermore, events such as the Tuskegee experiment (see Chapter 3) underscore why procedural transparency is so essential.[1]

5.2.3.2 Payments and Wrap-Up

At the end of the session, you should be sure to compensate the subject as specified. Compensation can take the form of monetary payment, credit toward a class, or nothing. If you are paying subjects monetarily, check with your supervisor, as there are nearly always detailed instructions for how to process such payments. In any case, you should make sure that they receive their compensation, that you receive any required documentation back from them, such as receipts or signatures, and that you thank each participant for his or her assistance. Without the participants, after all, you cannot run the study.

---

[1]The abuses associated with these studies led to the Belmont Report and the modern IRB process as a means of mitigating future risks to experimental participants.

At the end of the wrap-up, you should set up for the next subject. Make sure that copies of forms are on hand and that, if you have used such things as spare batteries, you restock.

5.2.3.3 Verifying Records

After each subject, it is a good idea to make sure data files are properly closed. For example, if an EPrime program is terminated not by running to its normal conclusion but by shutting down the computer, the data file may not be saved correctly. Any paperwork, whether it contains data (e.g., a questionnaire) or simply clerical work (how much credit should be given) should be verified and appropriately filed.

This is also an appropriate time to anonymize the data, as discussed in Chapter 3. You will, of course, want to retain a record of subjects' names for purposes of assigning credit or documenting payment, but if it is not necessary to associate subject names with data, the names should be removed as soon as possible. Depending on the nature of the research, you may want to store a list of subject codes and names that could later be used to relink identity information with the data, but you should consider carefully whether this is necessary.

It is also useful to keep notes about every subject. For example, if something unusual happened—the subject reported an apparent problem with the experimental software, the subject seemed to ignore instructions, a loud distraction occurred in the hallway—this should be noted so that the lead researcher or PI can make a judgment about whether to include that subject's data, whether to conduct additional tests on the software, and so on. Don't think, "I'll remember to mention this at the lab meeting"—you won't, at least some of the time. One of us asks our research assistants to initial a list of subjects to verify that everything went smoothly, including entering the correct information in the program running the experiment, starting on time, and so on. Sometimes, too, a subject will say something that provides an insight into the research question—if that happens, write it down at the end of the session. Such insights can be like dreams: clear and vivid in the moment and impossible to remember later.

It is also useful to document, perhaps in a lab notebook, information such as the date when particular data were collected (the dates on data files may reflect when they were last accessed rather than when they were collected), the file names for programs used to collect data, and so on.

This advice may seem obsessive, but it comes from long experience in running experiments. It is likely that the experiment you are running is one of many conducted in the laboratory you're working in, and perhaps one of

many that you're running yourself. Having a record you don't need is not a problem; lacking a record you do need may mean that the data collection effort was wasted or at least that you will need to spend a lot of time reconstructing exactly what you did.

5.2.4 Running Simulated Subjects

You may find yourself running simulated subjects. User models and simulations are increasingly used, both as stand-alone objects and sometimes as part of a study to provide a social context. For example, to model a social situation, you might have two intelligent agents act as confederates in a resource allocation game (Nerb, Spada, & Ernst, 1997). These agents provide a known social context in that their behavior is known and can be repeated, either exactly or according to a prescribed set of knowledge.

When you run simulations as subjects, you should keep good notes. There are often differences between the various versions of any simulation, and this should be noted. Simulations will also produce logs, and these logs should be stored as securely and accurately as subject logs. There may be more of them, so annotating them is very prudent.

If you create simulations, you should keep a copy of the simulation with the logs as a repeatable record of the results. You should perform enough runs that your predictions are stable (Ritter, Schoelles, Quigley, & Klein, 2011) and then not modify those files of model and runs but only copies of them.

Obviously, many of the issues discussed in this chapter do not apply to simulated subjects—no one, to our knowledge, has ever proposed that a simulated subject should be debriefed! Nevertheless, the importance of a clear protocol for your experiment is unchanged.

5.3 Other Issues

While running a session, issues may arise. Often these are negative, and you want to be prepared for them. Sometimes, however, you can learn something not anticipated.

5.3.1 Missing Subjects

Every study has two key parties—the experimenter and the subject or subjects (when running groups). Inevitably, you will encounter a situation where a participant does not show up, despite having an appointment. While participants should notify you in advance if they are going to be absent, keep in mind

that missed appointments do happen, and plan around this eventuality. Subjects are participating voluntarily (even when they receive compensation)—nobody is required to participate in a particular experimental procedure. Therefore, it is appropriate to be gracious about absences. Where possible, we recommend offering to reschedule once. However, when there are repeated absences, it is often not worth rescheduling. Bethel and Murphy (2010) estimated that about 20% of subjects will fail to arrive. This seems slightly high to us; for example, in the Psychology Department subject pool at our university, the no-show rate is typically 5% to 7%. In any case, the lesson is that you will have to schedule more subjects than your target to reach your target number of subjects, particularly for repeated session studies, studies with groups, or populations with understandable difficulties.

In some cases, you as an experimenter may need to cancel an experimental session. It is unacceptable for an experimenter simply not to show up for a session. When you really have to cancel the experiment, you should do it in advance. Furthermore, as the experimenter, you should cancel a session by directly contacting the participants.

Note that in some cases, there will be specific rules about these issues—for example, the policies of your subject pool may require 24-hour notice to cancel an experimental session or may have criteria for when absence is excused or unexcused. It is important to know and follow these rules.

5.3.2 Other Problems and How to Deal With Them

Most cognitive psychology and human–computer interaction studies run smoothly. However, if you run experiments long enough, you will encounter problems—software crashes, apparatus breaks, power goes out, and so on. Sometimes, too, there are more person-oriented problems—difficult subjects or problems that involve psychological or physical risks to the subject. Ideally, the research team will have discussed potential problems in advance and developed plans for handling them. It is the nature of problems, though, that they are sometimes unanticipated.

The most common problems are minor—software or equipment failures, issues with materials, and so on. In responding to such problems, the most important things to remember are (a) remain calm—it's only an experiment—and (b) try to resolve the problem in a way that does not cause difficulties for your subject. For example, computer problems are often solved by rebooting the computer—but if this happens 30 minutes into a 1-hour session and you would have to start over at the beginning, it is not reasonable to expect the subject to extend his or her appointment by half an hour, nor would the data

likely be usable. Often, the best thing to do is apologize, give the subject the compensation promised (after all, the subject made the effort to attend and the problem is not the subject's fault, and it is appropriate to be generous in these circumstances), make a note in the lab notebook, and try to fix things before the next subject appears.

It can be harder to deal with problems caused by difficult subjects. Sometimes, a subject may say, "This is too boring; I can't do this . . . ," or simply fail to follow instructions. Arguing with these subjects is both a waste of your time and unethical. As noted in Chapter 3, a basic implication of the voluntary participation is that a subject has the right to withdraw from a study at any time, for any reason, without penalty. Depending on the situation, it may be worthwhile to make one attempt to encourage cooperation—for example, saying, "I know it is repetitive, but that's what we have to do to study this question"—but don't push it. A difficult subject is unlikely to provide useful data, anyway, and the best thing is to end the session as gracefully as you can, note what went on, and discuss the events with the PI.

You can also encounter unexpected situations in which a participant is exposed to some risk of harm. For example, occasionally, a subject may react badly to an experimental manipulation such as a mood induction or the ingestion of caffeine or sugar. It is possible, though extremely rare, for apparatus to fail in ways that pose physical risks (e.g., if an electrical device malfunctions). And very rarely, an emergency situation not related to your experimental procedure can occur—for example, we know of instances in which subjects have fainted or had seizures while participating in experiments, and fire alarms can go off at any time.

Investigators must be committed to resolving these problems ethically, recognizing that the well-being of the participants supersedes the value of the study. If an emergency situation does arise, it is important that the experimenter remain calm and in control. If necessary, call for help. If the problem is related to the experimental procedure, it may be wise—or necessary—to cancel upcoming sessions until the research team has discussed ways to avoid such problems in the future.

It is important to bring problems and risks to the attention of the lead researcher or PI. In the event of problems that result in harm to subjects, it is important to consult the relevant unit responsible for supervising research, such as the IRB. These problems are called "adverse events" and must be reported to the IRB.

5.3.3 Chance for Insights

Gathering data can be tedious, but it can also be very useful. The process of interacting with subjects and collecting data gives you a chance to

observe aspects of behavior that are not usually recorded, such as the subjects' affect, posture, and emotional responses to the task. These observations that go beyond your formal data collection can provide useful insights into the behavior of interest. Talking informally with subjects after they have finished the experiment can also provide insights. You may find that subjects did not understand some aspect of the instructions. This is shown in Figure 5.4, where insights from debriefing may provide feedback (shown with the dashed arrow) that alters the setup and running of a session.

Obtaining these kinds of insights and the intuition that follows from these experiences is important for everyone, but gathering data is particularly important for young scientists. It gives them a chance to see how previous data have been collected and how studies work. Reading will not provide you this background or the insights associated with it; rather, this knowledge comes only from observing the similarities and differences that arise across multiple subjects in an experiment.

So be engaged as you run your study and perform the analysis. These experiences can be a source for later ideas, even if you are doing what appears to be a mundane task. In addition, being vigilant can reduce the number and severity of problems that you and the lead investigator encounter. Often, these problems may be due to changes in the instrument or changes due to external events. For example, current events may change word frequencies for a study on reading. Currently, words such as *bank*, *stocks*, and *mortgages* are very common, whereas these words were less prevalent a few years ago. Billy Joel's song "We Didn't Start the Fire" highlights these changes.

5.4 Running the Low Vision HCI Study

The example studies again illustrate concepts in this chapter. Judy's study illustrates the importance of piloting and what can be learned about the apparatus during the pilot study. While starting to set up the pilot study, Judy identified the experiment's first major issue: The company's software was not cross-system compatible; that is, it did not run on all versions of Windows. This was useful information and helped refine the experimental setup and protocol.

During the pilot study, the two pilot subjects (who were legally blind and not part of the subject pool) identified persistent text-to-voice issues. The team was able to successfully implement a version of the software that was cross-system compatible for the experiment, but the text-to-voice issues could not be entirely eliminated within the time period allotted for the study.

These problems caused Judy to reconsider her test groups, adding two additional groups. Besides the control group (unmarked navigation bar) and the first experimental condition (marked navigation bar), she added two other experimental conditions: (a) a customizable graphical interface controlled with the arrow keys, without a marked navigation bar and (b) a customizable graphical interface with a marked navigation bar.

The decision to add a customizable graphical interface was in response to the text-to-voice issues—the company's text-to-voice processing had a difficult time with book and movie titles, particularly if those titles included numbers. A major component of Judy's experiment tested the software's ability to support users' browsing book and movie titles. The relative lack of surrounding text in these lists caused the software's hidden Markov models to frequently misread years as numerals. Because the software's statistical tools for disambiguating between differing pronunciations also largely depended on surrounding text, Judy's text-to-voice software would in some cases mispronounce words—for instance, failing to distinguish between the noun and verb forms of the word *project*. Consequently, in the pilot study, Judy was uncertain if the lag times associated with the original experimental conditions were, in fact, a result of the treatment or confusion caused by the text-to-voice issues.

To isolate to some extent the effects associated with the software, Judy's team implemented a customizable graphical interface that allowed users to increase the size of a selected object with the up-and-down arrow keys and the color with the left-and-right arrow keys.

5.5 Running the Multilingual Fonts Study

In this example, developing our discussion from Chapter 4 regarding internal validity, we specifically discuss piloting to improve internal validity. Through piloting, we often find procedural or methodological mistakes that have consequences for an experiment's internal and external validity.

In the initial pilot data, Ying discovered a distribution in the data that she could not initially explain. The effect of changes in pixel density and size matched her expectations (denser letters were generally clearer, as were larger ones, with the magnitude of these effects eventually flattening off). Also as expected, she did find a relationship between matrix formats and these thresholds when the participants encountered a black font on a white background. However, she found that her color findings, even for Roman characters, did not match the literature. Previous work had shown that not only a font's size and density but also its brightness difference has an influence on its readability, and that light text on dark backgrounds and dark text on light backgrounds have

predictably different distributions. Ying and Edward's pilot data did not even remotely match the distributions found in the literature.

Ying, Edward, and the PI began brainstorming about the possible causes for this discrepancy. Looking through the pilot study's screening questionnaire, Edward noted that there were no questions regarding color-blindness. Further, the initial survey questions asked the participants to rank the matrix formats' colors relative to each other for formats of a given size and density. The initial list did avoid sequentially listing orange, red, and green matrix formats; however, it listed a blue matrix format followed by a yellow one. Many participants refused to complete the rankings because they could not see any distinguishable differences between the matrix format within a given size and density condition. Consequently, Ying's light-background/dark-font distribution was essentially bimodal and incomplete, where the bimodality was a result of whether the format was ranked or not.

To address this problem, Edward and Ying expanded the screening questionnaire to include questions about color-blindness. In addition, they replaced their relative ranking scale and replaced it with a Likert scale, where participants encountered each color for a given condition separately. They then could respond to the question, "Do you find this sentence easy to read?" by selecting one of five answers: strongly agree, agree, unsure, somewhat disagree, or disagree.

Summarizing the data required additional steps, because on this measure the relative emotional distance between selections cannot be assumed—the distance between *strongly agree* and *agree*, for instance, may be larger or smaller than that between *unsure* and *agree* for a given topic. So, for the purposes of summarizing the data, Ying had to group selections into positive and negative responses and then order the color format within a given pixel/density condition with respect to the number of positive or negative responses collected. Ying could then see the gradation in user preferences for the given brightness differences across the various matrix formats, both in the new pilot data and later in the study.

Piloting, in this case, led to a slightly different but better procedure. The data were slightly harder to analyze, but the data were cleaner, and color-blind subjects were not frustrated.

5.6 Running the HRI Study

The HRI (human–robot interface) study also illustrates issues that arise when running a study and resolution of those issues. A problem that Bob is very likely to find in running his study is that of recruiting suitable subjects.

Unlike universities, companies frequently do not have a lot of potential subjects available. Often, the only people readily available are those who know about the product or have a vested interest in seeing the product succeed commercially. These are not ideal subjects to test a robot. Bob will have to look into recruiting people through newspaper ads, casual contacts, and other contacts at and through the company.

When piloting his study, Bob found that the default video format for his camera did not match the video-editing tools. This was quickly and easily fixed by changing the video at the recording. Otherwise, it would have added a needless step. He also bought a large disk drive—two, actually: one to record to and one to store off-site—to back up his video recordings. (He was careful to encrypt the disk.)

In running his study in service of a company developing a product, Bob might find that he is tempted to terminate his study or controlled observation early when he finds useful results. Of all our examples, his study would be the most appropriate to terminate early because that is what he is looking for—changes that lead to a better product. He is not looking for a general answer to publish but is looking for results to improve his product. On the other hand, if the people he is trying to convince are skeptical, he may particularly want to finish the study because robots are hard to set up and maintain and more reports of subjects pounding the table in frustration may be more convincing. Similarly, if he finds dangerous conditions or results that are conclusive on an engineering level, he has an obligation to provide his feedback early and not put further subjects at risk.

5.7 Conclusion

Running the experiment is usually the culmination of a lot of work in developing the research question and hypotheses, planning the experiment, recruiting the subjects, and so on. It can also be the fun part, as you see your work coming to fruition and the data accumulate. There is a lot to attend to while running an experiment, but it is the last step before you have data to analyze and can find the answer to your research question.

5.8 Further Readings

Huck, S. W., & Sandler, H. M. (1979). *Rival hypotheses: Alternative interpretations of data based conclusions.* New York: Harper & Row.

This book provides a set of one-page case studies about how data can be interpreted and what alternative hypotheses might also explain the study's results. Following each case study is an explanation of what other very plausible rival hypotheses should be considered when interpreting the experiment's results. This book is engaging and teaches critical thinking skills for analyzing experimental data. It also reminds you of biases that can arise as you run studies. It would be useful as a reading for other chapters in the present text as well.

Mitchell, M. L., & Jolley, J. M. (2012). *Research design explained* (8th ed.). Belmont, CA: Wadsworth.

This book's appendix ("Online Practical Tips for Conducting an Ethical and Valid Study") offers useful tips similar to those in the present text.

5.9 Questions

Summary Questions

1. Describe the following terms.
   a. What is *debriefing*?
   b. List the procedures in debriefing, as noted by Mitchell and Jolly (2012).
   c. What is a *simulated subject*?
   d. What is an *adverse event*?
   e. What does *double-blind* refer to?
2. Consider where each of the example studies could be run in your environment, and draft a plan for doing so.
3. When can a subject decide to stop participating in a study?

Thought Questions

1. *Preparing a study*. Refer to the example scripts in Appendix 2. For your own study or one of the example studies, do the following:
   a. Prepare an experimental script.
   b. Prepare an informed consent form.
   c. Prepare a debriefing form.
2. Using only one figure or a similar, short (about 1-minute) task, prepare a short script to test the use of an ATM machine or logging into a computer. Run five people and note how you revised the script through using it.

3. Note how you would deal with the following potential problems in the study you are preparing or for one of the example studies:
   a. a subject becoming ill during the study,
   b. a subject getting lost and arriving 20 minutes late when another subject is scheduled to start in 10 minutes,
   c. a subject coming in an altered state,
   d. a subject self-disclosing that he or she has committed an illegal act on the way to the study,
   e. a subject disclosing orally his or her private medical history,
   f. a subject disclosing on a study form private medical history.