op 1

4 Data Recording

In the previous two chapters, we described the many aspects of eye-tracking systems (Chapter 2), and the importance of how to design an experiment to use in eye-tracker-based research (Chapter 3). This chapter takes us through the recording of the data.

The chapter is structured as follows:

- We begin with a hands-on summary in Section 4.1 (p. 111) to introduce the chapter.
- Section 4.2 (p. 111) discusses how to build the technical environment in which both the real data, and pilot data, are recorded.
- Section 4.3 (p. 115) covers the details of participant recruitment, and best practices when people arrive in your laboratory to take part in an eye-tracking experiment. It is important that participants feel good about the recording, not only to obtain good data and obey the laws of ethics, but also because you want participants or their friends to come back again and take part in future experiments.
- Section 4.4 (p. 116) deals with the eye camera set-up, and how to take on the challenges that some participants and environments pose to good data quality.
- Section 4.5 (p. 128) explains how to calibrate the eye-tracker on a participant.
- The instructions for the participants are important for acquiring valid data, and the most important points should be reiterated after calibration, and just before the start of recording, as Section 4.6 (p. 134) points out.
- In many cases, speech, galvanic skin response (GSR), electroencephalography (EEG), or other data sources are recorded together with eye-tracking data. As well as creating new possibilities, this also brings challenges, as described in Section 4.7 (p. 134) . This section may be regarded as optional if your sole purpose is to learn about eye tracking only.
- Debriefing (Section 4.8, p. 139) means telling the participant all he wants to know about the study. Commonly an information sheet is provided after the recording, and any additional questions are answered.
- Section 4.9 (p. 140) describes what to do once data are recorded; sorting and backing up files, as well as preliminary data analysis.
- The chapter is summarized in Section 4.10 (p. 143).

Building the recording environment and collecting data often starts before the experimental design is finalized. This is because piloting is such an important feedback mechanism to the formation of the experiment that many decisions cannot be taken before a pilot is made. When your real data are recorded, you have limited the number of possible experimental designs that you could previously choose from; the data produced relate precisely to the design you settled upon. Therefore, make certain exactly what you want to record before you record it. Carry out pilot recordings, and perform preliminary analyses of the pilot data.

Decisions made during the recording phase have a large effect on data quality, and so the researcher with the highest incentive to get good data should take part in the recordings, so she can influence the many choices made during eye camera set-up and calibration.

4.1 Hands-on advice for data recording

Recording data is the pivotal point in all eye-tracker-based research. All preparations lead to data recording, and all analyses use the data produced. Once you have started to use up your participants with a specific type of experimental set-up, you have committed to it.

- Test that any external equipment can interface with the eye-tracker correctly. Select stimulus presentation tools with care, and test them properly for timing.
- Be sure to *pilot* the technical set-up as well as instruction and data analysis properly before you bring in the real participants and start committing yourself to the recruitment and recording process.
- Treat participants so that they feel reasonably good about the task. Design the experiment so as to ensure that you get the most data and best data quality from the participants.
- Have all consent forms, rewards, and other documentation prepared and copied before participants arrive.
- Setting up the *eye camera* for a given participant is the most decisive factor for good data quality (aside from the eye-tracker itself), and requires an understanding of the principles of video-based eye tracking which can only be acquired by experience. The good thing is, this experience can be transferred to any other video-based eye-tracker.
- When he arrives, look at the participant's eyes to quickly identify any mascara, drooping eyelids, contact lenses, squint, or difficult glasses.
- Get acquainted with the degrees of freedom that you have in directing the eye camera towards the participant's eye(s), and learn how to use contrast, luminance, and focus settings to make it the best possible eye image.

4.2 Building the experiment

The preparation phase can be intense. Stimuli need to be prepared, the physical environment set up (unless the eye-tracker is fixed in a cubicle), and pilot testing needs to be run to make sure the experiment will actually work.

4.2.1 Stimulus preparation

Eye tracking is about looking at any type of stimulus, but the vast majority of all eye-tracking studies have stimuli in the form of still images, which may be simple or complex, shown successively one at a time. The order and timing of stimuli presentation is determined by the experimental design, and the stimuli are most often presented on a monitor. Images, perhaps `.jpeg` or `.bmp` files are shown one after the other according to the constraints of the design. More basic research into low-level oculomotor control and psychophysics often presents small dots or contrast patches on the screen, rather than full images. Other research fields, like psycholingustics, use images in conjunction with other presentation modes such as spoken utterances. Whichever field your research relates to, in most cases the stimuli are presented on a monitor with a fixed or controlled physical relation to the eye camera, and on- and offsets of stimuli are controlled by you and your stimuli presentation software. In some cases, eye tracking is used even though any visual stimuli are of no or peripheral interest. For example, tracking the eyes while presenting a narrative in a pitch-black room (Johansson *et al.*, 2006), or while stimulating the frontal eye fields with *trans-cranial magnetic stimulation* (Neggers *et al.*, 2007). Perhaps the most demanding stimuli preparation occurs when you use the real world as the recording environment with real objects as stimuli, as this involves physically constructing the stimuli, placing or moving them between trials, or post-hoc segmenting the recordings into actual trials. All these stimuli types have benefits and drawbacks that need to be considered by the researcher. An alternative to real-world stimuli is to trade some realism for better stimuli control. This would be the case for eye tracking in virtual environments, such as car simulators, where the world is like an interactive video.

Animations and videos are usually shown on monitors, and represent a step up in realism compared to still-image viewing, but add extra requirements on your technical set-up due to the dynamics. In particular, *synchronization* becomes critical. Be sure to check the following synchronization issues in your system:

- The timing between your stimulus program and the recording software of the eye-tracker: does the start of a new stimulus picture co-occur in time with the corresponding trial mark in the eye-tracking data?
- Are there cumulative synchronization errors between your stimulus program and the presentation of videos, due to variable frame rates caused by other processes competing for the CPU? If the trials are terminated immediately after video playback, does the same video consistently produce the same trial lengths?
- What is the timing between your stimulus program and other recording systems, such as EEG, GSR, and sound recordings?
- Do you assume a particular and fixed frame rate? Is the 24 fps video really 24 fps, or rather 23.976 fps (NTSC)?

In complex set-ups with many stimulus presentation forms and several recordings, synchronization issues often become difficult to solve, and you risk losing control over the important temporal dimensions, and end up with no or false results, as in Figure 2.22 on p. 46. Furthermore, videos (i.e. moving stimuli) are problematic due to their interaction with current oculomotor event detection processes, which may give you invalid data (p. 168).

However, in some research settings which use the real world as the stimulus, like real car driving, your participant looks at real-life traffic that you cannot control. Pedestrians come and go to their own liking, not according to your experimental plan. You simply have to take what you get. If you record from two people in a collaborative or face-to-face setting, you have the same lack of control. Participants may or may not say or do what you hope they will. In these cases, it is extra important to pilot your stimuli and your task, to make certain that you are likely to naturally encounter the behaviour that you want to measure.

If you are preparing a monitor-based experiment, there are a number of commercial stimuli programs that you can use: e.g. E-prime, Presentation, MatLab with Psychophysics Toolbox, PsyScope, as well as a number of dedicated slideshow presenters and experiment programs that have been developed in labs or by manufacturers. The commercial programs allow for more advanced experiments, and also for interfacing with additional soft- and hardware, such as audio, GSR (galvanic skin response), and EEG (electroencephalography), as well as mouse devices, keyboards, and button boxes. There are, however, important considerations when selecting what solution you want to use to present your stimuli.

Programming environments MatLab in conjunction with Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner, Brainard, & Pelli, 2007) and Python with PsychoPy (Peirce, 2007), are extremely versatile, but the learning thresholds keep some users away. Additionally, a programming approach often takes more time to implement and consequently it is usually only preferred if no ready-made tool can present the stimuli as desired. Some programming software may require you to manage licence numbers, licence servers, or hardware keys (dongles), which may cause problems if you often change stimulus computers and relocate your set-up. However, if you have the ambition to perform many experiments in the future, it will very likely pay off to incrementally implement your stimulus presentation needs in a cross-platform, open source programming environment. This will gradually reduce the time spent coding each experiment as you recycle a lot of code. Open source solutions means it is easy to share code with students and colleagues, as well as replicate others' set-ups. Coding skills are crucial in a laboratory working with large amounts of quantitative data, and having students or laboratory assistants work with modifying coded experiments could be a good learning opportunity for what may become future PhD students.

Professional stimuli presentation software Commercial programs, such as E-prime, Presentation, and PsyScope, provide a fairly high flexibility, significantly lower learning thresholds, and are quick to program using the graphical user interface. There is a large user community which means you can often get help to implement more advanced routines, and simpler experiments already have user-contributed templates. On the negative side, you are limited to the capabilities provided for you by the software developers, and you do not have access to the underlying code. Synchronization with the eye-tracker may be an issue, and there is also the question of handling licences and hardware keys. This is typically the most common option, as it offers a good balance between simplicity and capability.

Eye-tracker manufacturer software The quality of software varies between manufacturers. Overall, the programs are easy to integrate with the hardware as the companies use it themselves in demonstrating their equipment. No single manufacturer software supports all kinds of experiments, and manufacturer software is usually very limited, often just catering for their main customer segment. For instance, one may lack support for using the web as a stimulus, another may lack support for gaze-contingent research. There is no general advice as the softwares vary greatly, but always examine these programs sceptically for bugs, latencies, and odd data processing procedures.

Non-experiment programs Programs used for presentations such as Powerpoint, Impress, and Open Office Presentation should not be used for stimuli presentation if you care about latencies. They may be a quick fix to a slideshow experiment, but if you depend on correct timings, you run the risk of getting results you cannot rely on and, consequently, not be able to publish.

Real life On the positive side, using the real world will, naturally, have a high generalizability to real-world behaviour since the study is more natural. On the negative side, you get poor or no control over random events and noise inherent in a complex and real environment. Also, any synchronization with eye-tracking equipment will often be manual and as such you will not have clear-cut trial starts and stops. Consider a consumer approaching a target product shelf: when does the trial actually start? Will it be when the user is already in front of the shelf, possibly missing any pre-views of the shelf the consumer has had, or should it be as soon as the shelf is remotely visible from a large distance in the scene camera. Will the gaze cursor be meaningful from a large distance where you cannot make out what precise products the consumer is looking at?

All stimulus presentation software should be tested as part of building the recording environment. Do not accept them at face value.

4.2.2 Physically building the recording environment

Some experiments are simple enough so that you just use a permanent set-up, typically consisting of eye-tracker, recording computer(s) with appropriate software, monitor, keyboard, and mouse, possibly speakers, microphone, and video camera. In other cases, you have to build the environment from scratch; for instance if you are recording on paper newspapers with a head-mounted system and magnetic head-tracking, or if you are combining a virtual reality environment with eye tracking. Building the environment is the physical correspondent to programming the stimulus presentation.

With a static monocular eye-tracker and a stimulus monitor, set up the two so that the recorded eye is placed in the middle of the imagined volume extending from the monitor. In such an eye-centred set-up, you minimize the risk that saccades in the two horizontal extremes of the monitor will be different, although such placement effects should already be controlled for in the design by counterbalancing the stimuli positions.

Be careful with lighting conditions (p. 17), so that you do not have additional infrared reflections or poor contrast in the eye image. Be careful to have a stable environment to avoid environmental noise in your data (p. 35).

If you are using a system with magnetic headtracking, it needs to be set up separately—in a process that measures the geometry of the scene. Also beware of magnetic disturbances near the recording area, such as fans, elevators, and computer equipment of various kinds.

Similarly, magnetic systems such as TMS and fMRI may cause interference, as the magnetic field causes induction in the eye-tracker electronics. Therefore, most researchers working with video-based systems and TMS or fMRI prefer to use remote or even long-range systems.

4.2.3 Pilot testing the experiment

Pilot testing the experiment is essential. Always pilot! Everything will improve, in particular your stress level during real data collection. There are a multitude of reasons why you should pilot your experiment:

- You find out what cables are missing in the eye-tracker and stimulus set-up.
- You will be able to see if your proposed stimuli and instructions do not elicit the behaviour that you want to measure and quantify, so you can get to change them well before the real recording.
- You practise handling of participants and equipment for this particular experiment.
- You check the geometry of the recording environment, and that you can calibrate and record data in all gaze directions required by your experiment.
- You will find out whether the software that you thought would work in fact does not.
- You can ask your pilot participants how well they understood your task instructions.
- You will see whether the proper data files are generated by the eye-tracking and stimulus software and that these contain the data you will be using.
- You can run a pilot analysis on the pilot data to see that the selected dependent variables and your method of analysis really work.
- Running a quick pilot or dummy recording at the beginning of each day that you will be recording will help you discover if your set-up has been tampered with; perhaps somebody borrowed a cable or some other critical part and forgot to return it or reconnect it properly?

After piloting, alter your experiment so that it does what you want it to do, and then reiterate the pilot. When you are happy with your pilot study, you can start recruiting the real participants.

4.3 Participant recruitment and ethics

When booking a participant for your study, allow time for him to ask questions. Prepare what you will answer when the participants ask questions, which they will. The most common question tends to be "What is the purpose of your study?". Have a good cover story, or tell them that they will get to know afterwards. Prepare a list so that you can offer to send participants the general results, your publications, etc. Other difficult questions that sometimes come up are "Is this testing my intellectual abilities?" and "Can I do anything wrong?"

Decide beforehand who will be in the room during the experiment. In some studies, participants will feel more comfortable if they are left alone with their task.

When he arrives, welcome the participant and introduce him to the equipment and to the people working with the experiment. Be social to him. At the same time, be certain to look at his eyes. Look for things that may mess up your recording: mascara, downward pointing eyelashes, certain types of glasses, and contact lenses. Is the participant squinting? The sooner you know about such problems, the better you can plan what to do. This is particularly true if you are on a tight recording schedule.

Also, remember that the second a participant shows up for your experiment, his thoughts are occupied mostly with non-laboratory ideas; home-work assignments, evening schedules, romantic plans and other pressing business. Different participants will be in different states of mind, and they will associate your instructions with slightly different things depending on their reference frame of the world and of themselves. Similarly, their physiological states will differ. Some have come running directly from a class, their minds working at full capacity with all newly learned ideas, and others have come directly from home, still not wide awake. These differences will contribute to the heterogeneity of the participant sample and may add variance to data, which may detrimentally influence your later statistical analysis.

One thing the participants *do* have in common is that they are all located in your laboratory for testing, and this you do have control over. You should give them a chance to settle and assume a similar state of mind. Allow the people coming directly from class to sit down and gather their thoughts, and allow the people coming straight from home to take a minute to wake up and think about why they are here. Let the laboratory setting work its magic and frame the procedure for each participant in the same way as you take them into the experiment room and give them their instructions. Keeping these things constant makes it more likely your instructions will be interpreted similarly across participants, and this will help you to get a more homogeneous sample.

If your study allows it, showing the eye camera image to the participant as part of setting it up may help the participant to cooperate better and move about less. If your participant is a child, showing and explaining what you are doing is particularly important, since a child is more likely to move his eye out of the eye camera view. When you do work close to the participant, such as adjusting the equipment, tell him in advance exactly what you are going to do, so you do not startle them. Be sure to tell the participant that it is OK to stop the experiment at any time. You do not want your participant to silently suffer through the second half of your experiment.

4.3.1 Ethics toward participants

Ethics towards participants are comprised of two elements: 1) The law, which varies across regions in the world. 2) The researcher's common sense, empathy, and desire to have a good relationship with the participant pool.

In many parts of the world, experiments on children, clinical populations, or participants that are likely to experience some degree of psychological stress due to the task, require the experimenter to apply to a regional board of ethics before running the study. The ethical board will then decide whether the study follows the local law, and whether the scientific purpose of the study outweighs the hardships and/or sufferings of the participants. You may think that participants do not suffer very much from taking part in an eye-tracking study, but participants with clinical diagnoses, such as dyslexia, can feel bad about doing a reading study in a highly technical laboratory, in particular if the participant is young and has already been subjected to excessive testing.

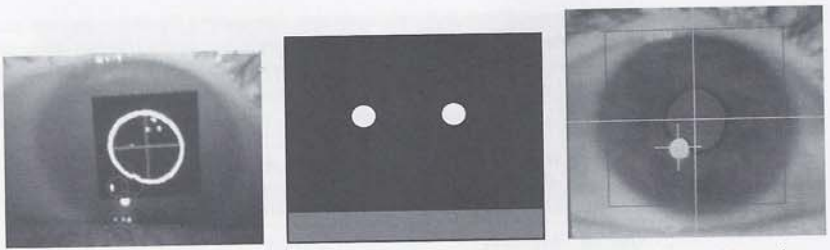
Ethics fundamentally means that the participant should have a good feeling about having participated in your study. This is in your own interest, because your participant will tell his friends about your experiment and the laboratory in which it took place. Not only will you increase the chances that your participant will come back therefore, you potentially gain new participants for future studies also. If you have treated your participants well, and if they are interested in your study, they may even recruit their friends for you! If you get them to be really interested, after a few years, one of your former participants may suddenly be your student. All in all, you should treat your participants well.

Many ethics-related issues are dealt with in the 'consent form', a written statement that the researcher may use the data collected from the participant, and that the participant considers himself fully informed about the purpose of the study and the consequences for himself of signing off the rights to the data. Laws may require consent forms to be signed *before* data recording, but the objectives of the study sometimes require this to be done *after* the recording (so that the participant is not aware of the purpose of the study during the recording). This can be solved by preparing two consent forms: one which is more general in its experiment description, and a second post-experiment consent form which fully explains the experiment and gives the participant the option to have the recorded data destroyed if no consent is given.

4.4 Eye camera set-up

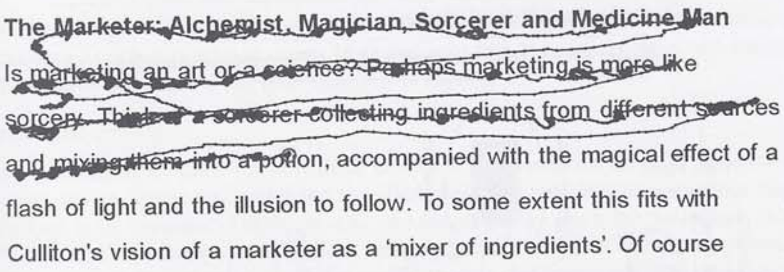
Now we are ready to put the participant into one of the eye-trackers described in Chapter 2. This section describes how to adjust the eye camera view so as to get data of high quality. Eye camera set-up is of great importance to the quality of your data, and largely decides whether you can use them in analysis, and for what measures. Knowing your eye video and the algorithms that calculate gaze position from it allows you to get better data from a larger spectrum of participants, more quickly and with less anxiety.

Some manufacturers have hidden the eye camera set-up in automatic processes, giving an appearance of increased usability and simplicity (Figure 4.1(b)), but at the same time withholding control over the data recording from the user of the system. The result in terms of data quality may be acceptable in a large number of cases, but some loss of quality will go unnoticed, or be difficult to understand and alleviate. Experienced users of these mostly remote systems will with time learn to move or tilt a whole system (camera unit and monitor) to set up an optimal eye-camera angle, behaving in accordance with the advice we give below, but without the feedback that a full eye-image can provide. Most eye-tracking recording software shows one or another image of the participant's eye, however. Figures 4.1(a) and (c) are examples from the ASL MobileEye, and the SR EyeLink respectively. In this section, we will use examples from SMI and EyeLink systems, because the video image can be very clearly seen with their systems, but the presentation and majority of advice in this section are valid for all video-based pupil–corneal reflection systems. We will use reading data to show the effect on data quality, because the lines of text provide a good reference point against which to compare data.



(a) The ASL MobileEye eye image (b) The Tobii Studio eye image (c) The EyeLink 1000 eye image

Fig. 4.1 The eye images of three video-based pupil-corneal-reflection eye-tracking systems. To the left ASL MobileEye eye image, in which the highlighted ring indicates the circumference of the pupil. In the middle, Tobii Studio in which the eye image is hidden, and to the right the EyeLink eye image, with pupil and corneal reflections clearly marked by colour and cross-hairs.



The Marketer: Alchemist, Magician, Sorcerer and Medicine Man

Is marketing an art or a science? Perhaps marketing is more like sorcery. Think of a sorcerer collecting ingredients from different sources and mixing them into a potion, accompanied with the magical effect of a flash of light and the illusion to follow. To some extent this fits with Culliton's vision of a marketer as a 'mixer of ingredients'. Of course

Fig. 4.2 Example of good quality data in raw format, recorded with a tower-mounted eye-tracker at 1250 Hz. The figure shows raw gaze samples; one dot per sample. Fixations are seen as blobs of many dots in a small area. Saccades are strings of dots; the more sparsely the dots are placed, the faster the saccade.

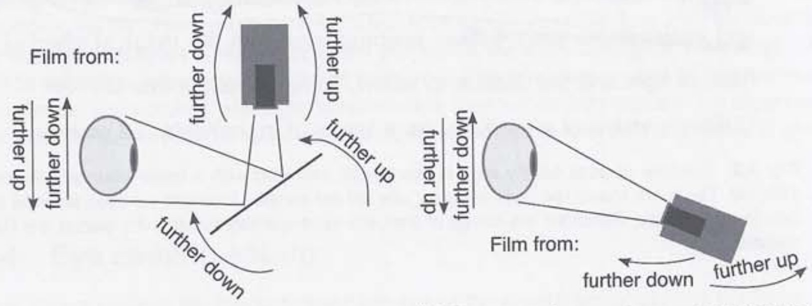
Errors can arise in the eye (feature) detection and gaze estimation algorithms at the heart of the eye-tracker due to *covering* of the pupil or corneal reflection, *confusion* between the pupil or corneal reflection and a number of other optic entities, *distortion* of the image, or because of *loss* either of the pupil or corneal reflection by the camera and its supporting software. A number of causes behind these problematic conditions are summarized in Table 4.1.

In order to get high-quality data such as that shown in Figure 4.2, you should see to it that the eye camera records the eye from slightly below the eye with regard to the direction the participant has when facing the stimulus area. If your eye camera is too low, it will be difficult to calibrate and record in the upper corners of the calibration area because the corneal reflection is more easily occluded by the lower eyelid. If the camera angle is too high, the upper eyelid will often cover some participants' pupils when looking at the bottom corners, and calibration will be difficult or faulty. Figure 4.3 shows how camera positions, the participant, and the mirror should be moved to achieve a better image. When you are moving cameras or mirrors close to your participant's eyes, take care always to talk to him about what you are doing, so that he is prepared and you do not startle him. Remember that by participating, he puts trust in you. Of course, you must be careful not to hurt him or scratch his glasses.

The participant and cameras should be placed in a position so that the eye camera shows a good image of the eye for each corner of the calibration area. This involves directing cameras and/or mirrors as well as raising or lowering the participant's chair, or the table with eye-

Table 4.1 Summary of optic conditions that may endanger data quality. For instance, both pupil and corneal reflection may be covered; the pupil foremost by droopy eyelids, and the reflection by laughter, or more specifically, the accompanied narrowing or closure of the eye. Confusion refers to cases when other objects are mistaken for the whole or part of the pupil or corneal reflection. Optic distortion can be caused by, for instance, bifocal glasses, while extreme gaze angles may cause loss of the corneal reflection.

	Pupil	**Reflection**
Covering	Droopy eyelids	Laughter
Confusion	Mascara Glasses Ambient infrared light Retinal reflection Specks and dirt	Retinal Glasses Contact lenses Sunlight Lamps Other infrared light sources Wet eye
Distortion	Bifocal glasses	
Loss	Head movement	Extreme gaze angles



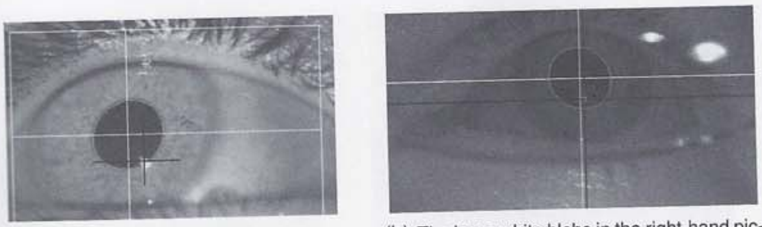
(a) Tower-mounted and some head-mounted systems

(b) Remote systems and some head-mounted systems

Fig. 4.3 How to move the eye camera, the mirror, or the participant, so as to get a camera angle from further up or further below.

tracker and stimulus monitor. If you are using a head-mounted system, camera set-up often involves positioning and directing a whole construction of mirror and cameras. If you are using a remote eye-tracker or a head-restraining static system, remember that participants tend to sit very attentively during camera set-up and calibration, but may slide down during the actual recording, so your camera angle changes or you have to move mirrors or camera during recording (or it is done automatically), which may cause an offset in data.

A pupil–corneal reflection system calculates the participant's gaze direction on the basis of the relative position of the pupil and the corneal reflection centres. When calculating the centre of the pupil, it is important that the eye camera sees the entire pupil in all gaze directions. Two examples of good eye images can be seen in Figure 4.4. Partial occlusion of the pupil causes a movement of the calculated centre point, and thus a movement in the data coordinates, although no real eye movement has been made. In fact, even very small movements of the calculated pupil centre can cause considerable movement in the coordinate files. The identified corneal reflection may also move artificially, jumping to other reflections



(a) Both pupil and corneal reflection are uniquely identified with white and dark cross-hairs, respectively.

(b) The large white blobs in the right-hand picture are infrared reflections in the participant's glasses, but they do not interfere with the gaze estimation algorithm.

Fig. 4.4 Eyes with good eye camera set-up, filming from below. Both eye images are from head-mounted systems.

in the eye video, most commonly just under or on top of the upper eyelid. This movement of the calculated corneal reflection causes large and very fast, but entirely false, movements in the data samples, which we call *optic artefacts* in Chapter 2.

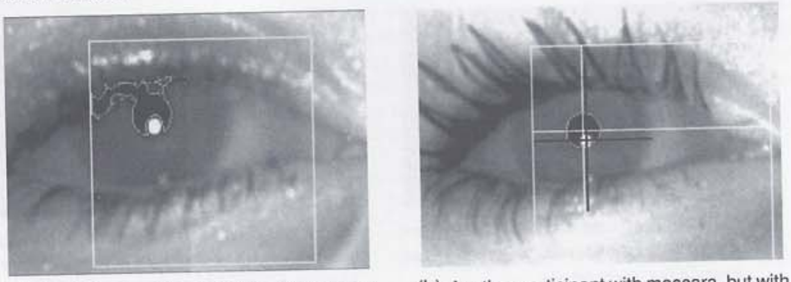
Ocular dominance

If your participant is squinting or if there is reason to believe that one eye is better than the other, see to it that you select the so-called *dominant* eye. Ocular dominance can be determined using a number of different tests; one simple test for this is the 'Miles test' (Miles, 1929). Here, the participant extends both arms and brings both hands together to create a small opening while looking at an object with both eyes simultaneously. The observer then alternately closes the left and the right eye to determine which eye is really viewing the object. That is the eye that should be measured. Around two thirds of the population have a right eye dominance and men are more commonly right dominant (Eser, Durrie, Schwendeman, & Stahl, 2008).

If you are recording monocularly and you notice that you have to change the camera to the other eye, for instance recording a left eye after a series of right eyes, you may need to take care that the position of the left eye relative to the monitor will be the same as that of the previously recorded right eyes (if your experiment demands control over this). If not, the result may be small variations in fixation and saccade data in the outskirts of the monitor. The best solution is to control for this in the experimental design, so horizontal stimuli positions are counterbalanced.

4.4.1 Mascara

Mascara is considered a serious problem for data quality. This may be less of an issue for a bright-pupil system, as the pupil is bright but the mascara dark (see Applied Science Laboratories, 2011), but mascara also blocks filming through sparse eyelashes, making it not completely unproblematic for bright-pupil systems. Therefore, regardless of eye-tracking system, always tell your participants beforehand not to wear mascara. Do not tell them not to wear make-up, because some participants think that mascara does not qualify as make-up. If the mascara is left on, only expect to make a good recording on participants with large, open eyes and upward eyelashes. This is because in a dark-pupil system, the software that identifies the pupil is confused by the other large dark area in the immediate vicinity of the pupil, and may lock onto the mascara rather than the pupil. This is detrimental to subsequent calibration, as well as to the recording. A make-up removal kit should therefore be an indispensible piece of



(a) Participant with drooping eyelid and downward eyelashes. A thick, dark brush of lashes melds with the pupil and makes it impossible for the recording software to identify the pupil.

(b) Another participant with mascara, but with upward-pointing lashes; in this case the mascara is easy to exclude, and lashes like these seldom occlude the pupil.

Fig. 4.5 Mascara interferes with pupil detection.

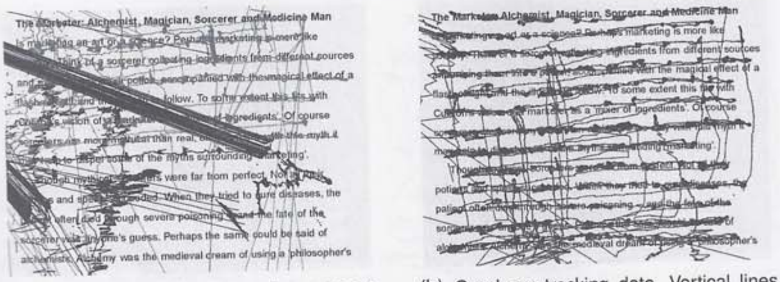
equipment in all serious eye-tracking laboratories. If you ask your participant to remove the mascara, watch out for any remaining mascara blobs that often have a round pupil-like form which can be mistaken for a pupil and later cause offsets in your data.

Some recording software can exclude portions of the eye video from the eye feature detection process, which allows quality recordings without asking participants to remove mascara. In Figure 4.5, only the parts of the eye video that are inside the white rectangle are analysed for pupil and corneal reflection. This works relatively well in the majority of cases, provided the lashes point upwards and the head is fixed so the eye does not move relative to the eye image (as for a fixed contact system with chin rest). Moreover, software programs often give the user the option to reject pupil-like objects that are too small or too large, such that mascara blobs with extreme sizes can be discarded automatically.

The right side of Figure 4.6 shows a raw, unfiltered plot of data samples from a participant for whom we had a good eye video in all gaze directions. It shows one pixel for each sample, of which there are 1250 per second. Where samples cluster to blobs, the eye has been still in a fixation. Where samples trace a line, the eye has moved quickly in a saccade. The vertical pairs of lines are blinks; the eyelid going down and then up again. All else is noise of various kinds. It can been seen that we have fixations and saccades almost right on the word for each line. The left side of Figure 4.6 shows data from a participant with mascara and downward eyelashes. The data plot is largely OK all the way to line eight, which is the gaze direction at which the mascara aligns with the pupil in the eye video. Data for the lower five lines consist of useless optic artefacts.

4.4.2 Droopy eyelids and downward eyelashes

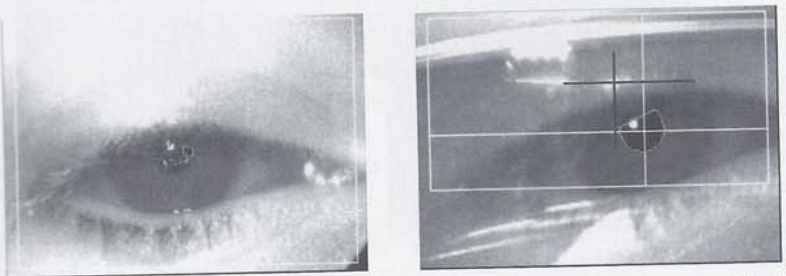
The problem with droopy eyelids grows with the age of your participants, but it is also a matter of individual variation. Droopy eyelids are a major problem in eye-tracking research even if there is no mascara, because the eyelids, or the eyelashes, cover the pupil in the lower gaze directions (Figures 4.7 and 4.8). In your data, you will see very large downward offsets, because a pupil partially occluded from above has a lower mass centre and resulting data will be at an artificially lower vertical position. At a certain gaze angle, the pupil is completely covered, and you will then have complete data loss in the lower part of the screen/visual field. Notice that pupil occludance may not appear dangerous during calibration, when the participant is more tense and focused, and the eyes are more open. It is when he gets into



(a) Data from participant wearing mascara and having slightly downward eyelashes. Data plotted in the upper half of the stimulus image shows data recorded when the eye was open and all of the pupil seen. Below this is the data where the pupil is covered by lashes with mascara.

(b) Good eye-tracking data. Vertical lines are blinks. Some noise can be seen in the upper left corner and the bottom line. However, all lines of text have accurate and precise data samples on top of them, that the event detection algorithms can make use of.

Fig. 4.6 The effects of mascara on slightly downward eyelashes.



(a) A thick brush of lashes covers the pupil and make it impossible for the recording software to identify either pupil, or corneal reflection.

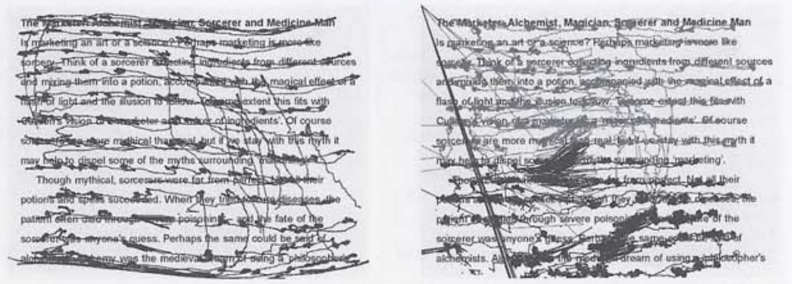
(b) In this drooping eyelid, on a participant with glasses, the eye-tracker locks onto a false corneal reflection on the eyelid, because the real corneal reflection is partially occluded underneath lashes. Also notice the partial pupil occlusion. Gaze data from this eye video will have considerable data loss, a very large offset, and many optic artefacts.

Fig. 4.7 Participant with drooping eyelid and downward eyelashes.

the task and relaxes that the eye closes and you get an offset. In fact, you may get better data quality if you calibrate the bottom calibration points with a relaxed participant, who closes his eyes a little, because that is what the eye will look like when you record your participant looking in that direction. However, that is a gamble and the correct solution is to redo the set-up to handle the drooping eyelids or the eyelashes.

There are at least four possible solutions for droopy eyelids and downwards eyelashes:

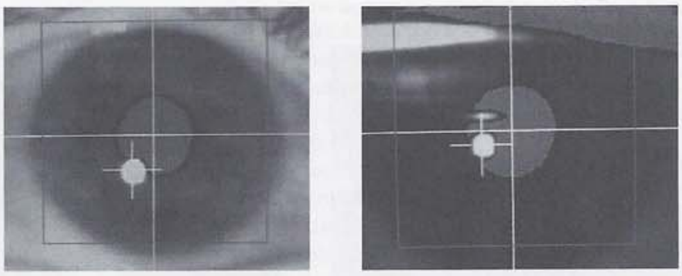
1. If possible, move the camera or the mirrors to film the eyes from even further below.
2. Ask the participant to use an eyelash curler to turn the eyelashes upwards. This instrument should be desinfected between participants.
3. Is the participant tired? Ask him to return at another date and time that is more appropriate.



(a) Mild effects. The bottom five lines have a one-line offset, and the two last lines are smeared together into one thick data line.

(b) Strong effect. Note the offsets and optic artefacts in the bottom four lines, and the centre. The data in the upper paragraph are fairly good, since the eye is open while the participant is looking there.

Fig. 4.8 Example of the effect in reading data of droopy eyelids.



(a) No glasses

(b) With glasses

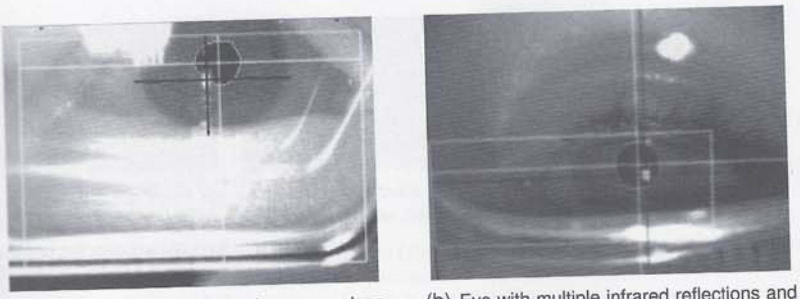
Fig. 4.9 Effect of glasses on one and the same participant. Note in (b) the extra reflection just above the corneal reflection, the darker image with lower contrast, and the large reflection in the upper left corner.

4. As a final resort, use some sticking plaster to fixate the participants upper eyelid. This works well for most elderly participants, but may not be comfortable for some participants.

4.4.3 Shadows and infrared reflections in glasses

Glasses make eye tracking more difficult in several ways. First, glasses, or possibly the surface treatments of the glasses, may make the eye image darker, reducing the contrast between pupil and iris, which may decrease accuracy and precision in a dark-pupil system. Access to contrast and brightness settings in the eye camera is useful to remedy this. Second, the light from the corneal reflection can be reflected back to the eye and give a second, but fainter corneal reflection higher up in the eye image (compare the eye images in Figure 4.9). Third, the shadows from the brim may confuse the pupil detection in dark-pupil systems. All these three effects are seen in Figure 4.10.

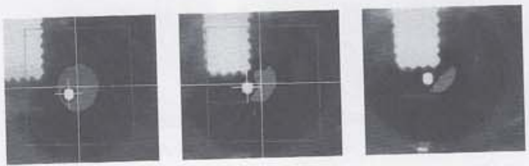
Another major problem is infrared reflections in the glasses themselves. If the infrared reflection is on top of the pupil or near the corneal reflection in any of the gaze directions—ask your participant to look around—your calibration or your data recording will be jeopardized. This problem is particularly large if the participant wears old, scratched glasses, or glasses



(a) Eye with shadow from glasses, and an infrared reflection high above the eye. Dark areas could confound the pupil detection but have been excluded from detection.

(b) Eye with multiple infrared reflections and refractions from glasses, both above and below the pupil. Were the eye to move, the pupil might be covered.

Fig. 4.10 Shadows and reflections in glasses.



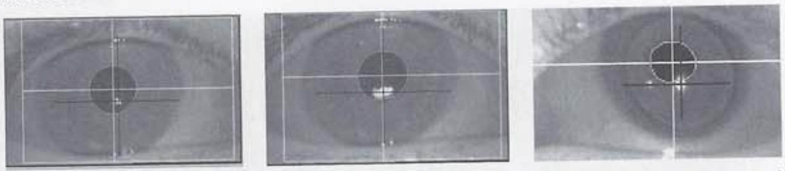
(a) Tracking works (b) Tracking with offset (c) Tracking lost

Fig. 4.11 Reflection in glasses covering more and more of pupil and finally making tracking impossible.

that have been treated to reflect sunlight, while glasses with anti-reflection treatment tend to elicit much fewer reflections. The solution is to move the eye camera angle so these reflections appear far from the pupil and corneal reflection (see Figure 4.11). It may take a while to find a view on the eye without interfering reflections, in particular with tower-mounted eye-trackers, but then these systems usually have good tracking throughout the experiment as the participant is fixed. For remote eye-trackers you are usually limited to changing the filming angle and then hoping that the participant does not switch to a disadvantageous position during the experiment.

Some people wear small glasses with thick designer frames. This may cause a dark shadow on part of the eye, which can interfere with both pupil and corneal reflection. Again the solution is to move the camera position and angle until an optimal viewing angle is found.

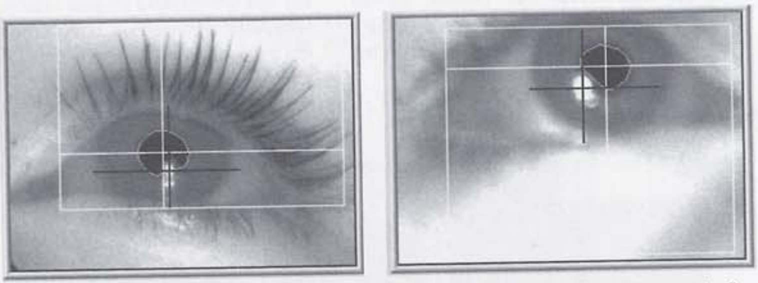
Some glasses are more difficult than others, in particular if they are so dark that the contrast in the eye video is too low, very small so the frame comes too close to the eye, or because they are very scratched and create many infrared reflections. A frame too close to the eye is problematic because it may occlude pupil or corneal reflection in certain eye positions, or the frame itself may be mistaken for a pupil or a corneal reflection if it is dark or reflective. If your eye-tracker allows you to change the luminance (and contrast) of the eye image, do that. Another solution, used by some eye-tracking researchers, is to have a set of their own eye-tracker-friendly glasses of different strengths that they lend to participants with problematic glasses.



(a) No air bubbles at the point of the corneal reflection.

(b) The effect of air bubbles on the corneal reflection.

(c) Hard contact lens; no air bubbles, but if the border of the lens passes over the corneal reflection, a similar offset will occur.



(d) Focused camera. Two corneal reflections; cross-hair on the upper one.

(e) Unfocused camera. Several corneal reflections merged into one, so tracking is stable, but the pupil detection is deformed, giving a small offset.

Fig. 4.12 *Soft contact lenses* in (a), (b), (d), and (e): Reduce focus on camera to avoid multiple corneal reflections. The *hard contact lens* in (c) does not generate bubbles.

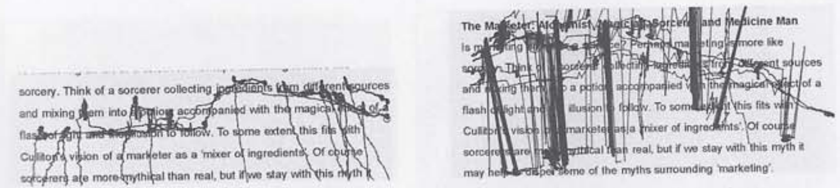
4.4.4 Bi-focal glasses

Bi-focal glasses introduce a border in the midst of the eye video that makes calibration difficult and recording close to impossible. Ask your participant to wear other glasses.

4.4.5 Contact lenses

Soft contact lenses cause only one problem in eye tracking, but it is a major problem, and it is not uncommon. For some people, because of a less than perfect fit between lens and eyeball, small air bubbles gather underneath the contact lens. When the infrared illumination is reflected in such a collection of small air bubbles, the light is split up into a number of reflections (Figures 4.12(b) and 4.12(d)). As the eye moves, the bubbles shift. Your eye-tracker will randomly select any of these reflections as the corneal reflection, and jump between them. This results in erroneous data samples and apparently very fast movements of the eye (Figure 4.13), known as optic artefacts. When this occurs it can be devastating for your data quality. You cannot always see the bubbles while setting up and calibrating a participant; they may appear in the midst of data collection, ruining your recording session.

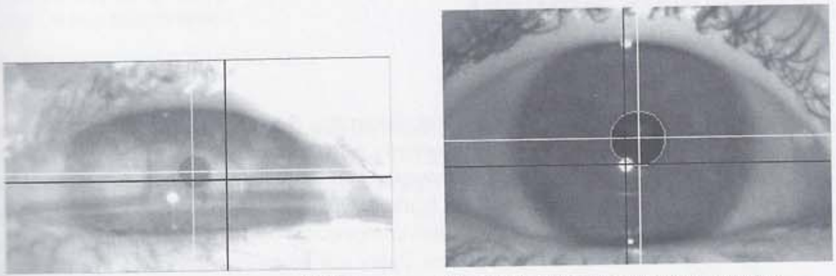
Fortunately the solution is mostly very easy. If your participant wears soft contact lenses, which you should ask them about or can even see in the eye image, always reduce the focus of the eye camera somewhat. This way the many small bubbles will merge into one larger corneal reflection. Your data will sometimes be slightly less accurate, because the larger corneal reflection pushes the pupil slightly to the side. Nevertheless, an unfocused recording is much, much better than if you had continued with full focus and a split corneal reflection.



(a) Participant reading one line of text.

(b) Participant reading two lines of text.

Fig. 4.13 Offsets and optic artefacts, due to contact lenses causing a split corneal reflection. The length of the false movement lines corresponds to the distance in the eye video between the two corneal reflections. As Figure 5.14 on page 163 shows, velocities in these artefactual movements is far beyond that of saccades.



(a) Full sunlight. Low contrast and high luminance; pupil is identified but corneal reflection is not.

(b) Indoors with only artificial illumination.

Fig. 4.14 Same participant wearing a head-mounted eye-tracker.

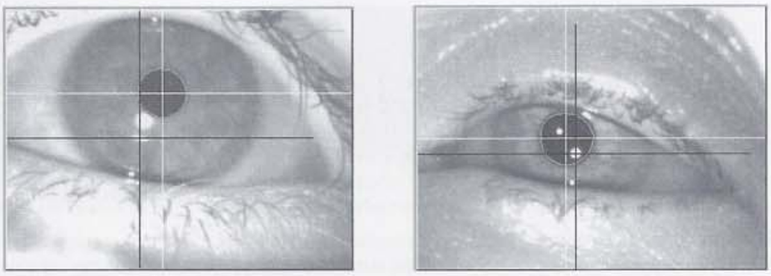
This will work in all cases, except the most extreme and rare ones, when the corneal reflection splits into a multitude of points that can cover up to a quarter of the iris.

Hard contact lenses are formed to fit the eyeball much tighter, and appear to gather no air bubbles, and data can be recorded at full eye camera focus. Hard contact lenses are very clearly seen in the eye camera image, as in Figure 4.12(c). Only if the border of the hard lens moves across the corneal reflection, which can occur often for some participants, will data quality suffer.

4.4.6 Direct sunlight and other infrared sources

Direct sunlight contains much infrared light, enough to outshine the infrared illumination many times over. The result is typically complete and immediate data loss (compare Figure 4.14). This can be particularly difficult during car driving, especially when the sun is close to the horizon, shining directly onto the eyeball. One of the authors was once in a car-driving project where we drove along the Scanian coast with a western sun shining through a sparse forest. The resulting rapid alternations between sun and shadow yielded equally rapid alterations between data capture and data loss. Such data, of course, are useless. In cars, one option is to attach an infrared filter film to the windows, but a more common option is to simply record on a cloudy day. The SmartEye systems that are specifically designed for use in cars select a part of the infrared frequency spectrum in which the sun has a lower effect on the eye-tracker camera, compared to the normally used infrared frequency.

Static systems can also be affected by sunlight, for instance through a window. Good shades or no windows at all are therefore recommended in an eye-tracking laboratory.



(a) A hot incandescent bulb gives a second reflection in the cornea, and the corneal reflection cross-hair has locked onto it.

(b) The fast-moving fourth Purkinje reflection in the upper left quarter of the pupil.

Fig. 4.15 Additional reflections.

A video-based motion capture system illuminates the scene to be measured with several infrared lamps, which could seriously compete with the infrared diode of the eye-tracker. We have on several occasions tested head-mounted systems with motion capture systems such as Qualisys. This combination has always worked. Obviously the shorter distance from the eye to the eye-tracker infrared diode more than compensates for its weaker luminousity.

In fact, some ceiling-mounted indoor lights can be more harming to eye tracking than motion trackers as measured by the intensity of the corneal reflection that they give rise to. In Figure 4.15(a), an incandescent light from the ceiling caused an additional corneal reflection, which is similar to the double reflection of the contact lenses, but only now it is permanent. If we were to calibrate on such an eye image, the danger is that some of the calibration points are calibrated with the correct reflection, and other on the false reflection, as in Figure 4.15(a). Terrible offsets result from such a calibration. Solutions include turning off the light, or using a lamp that does not emit infrared light,[14] and if that is not possible, use recording software settings to select properties such as reflection perimeter and distance to pupil centre to tell the eye-tracker which reflection to use.

4.4.7 The fourth Purkinje reflection

You may sometimes see an additional reflection inside the pupil, as in Figure 4.15(b). It moves very quickly when the participant moves his eyes, but as soon as it reaches the border of the pupil, it disappears. It is produced at the back of the lens, and can therefore only be seen through the pupil. This reflection is used by dual-Purkinje eye-trackers, and is the reason they are so precise and have such a small tracking range. It may on rare occasions—for only a few samples—interfere with the corneal reflection and undermine data, but it is the weaker one and will not overly displace the cross-hair from the corneal reflection, as long as the eye camera is focused. If the camera is out of focus, this reflection may be much bigger and undermine the calculation of the pupil area.

4.4.8 Wet eyes due to tears or allergic reactions

When the participant's eye is wet, the corneal reflection splits up into several reflections, similar to the split corneal reflection in a contact lens. This may happen due to allergic reactions in the pollen season, or if you instruct the participant to try not to blink. The latter case makes

---

[14] You can check any lighting source for infrared by covering the eye-tracker's infrared source and pointing the camera at the light—if the system picks it up you know it emits infrared light.

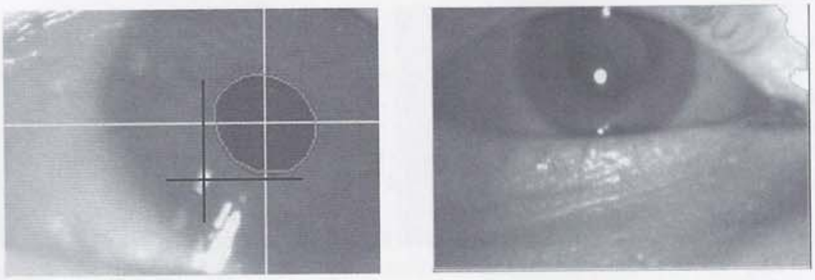
urth Purkinje reflection
r of the pupil.

e measured with several
de of the eye-tracker. We
capture systems such as
er distance from the eye
er luminousity.
ing to eye tracking than
n that they give rise to. In
tional corneal reflection,
nly now it is permanent.
e of the calibration points
tion, as in Figure 4.15(a).
ing off the light, or using
e, use recording software
to pupil centre to tell the

as in Figure 4.15(b). It
n as it reaches the border
nd can therefore only be
rackers, and is the reason
rare occasions—for only
data, but it is the weaker
reflection, as long as the
may be much bigger and

o several reflections, sim-
n due to allergic reactions
nk. The latter case makes

nfrared source and pointing the



(a) Allergic reaction resulting in a wet eye and multiple corneal reflections. The identification cross is on a false reflection.

(b) Infrared reflections in the retina giving a semi-bright pupil.

Fig. 4.16 Reflections in tears and retina.

them gaze until their eyes are dry and then the eye compensates with more tear fluid. The distance between reflections is often larger in this case, so the problem cannot be solved by decreasing eye camera focus. However, it mainly appears when gaze is in the far upper corners, as in Figure 4.16(a), and can mostly be solved during manual calibration with little or no data loss.

4.4.9 The retinal reflection (bright-pupil condition)

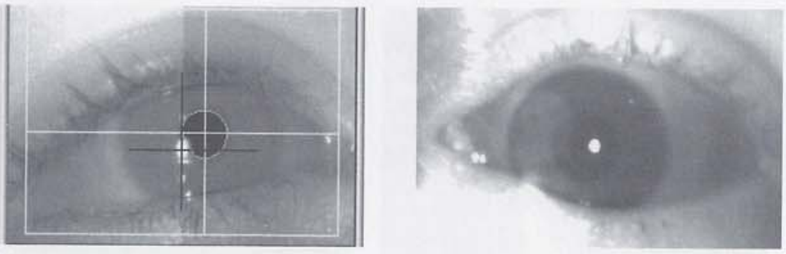
On rare occasions in dark-pupil systems, you may happen to have the infrared illumination almost co-axial with the line of sight, and then see the reflection in the retina. Unless you are using a bright-pupil system, you do not want this, because it makes pupil identification difficult or impossible, as in Figure 4.16(b). This condition can be solved by moving illumination and the mirror.

4.4.10 Mirror orientation and dirty mirrors

Wrong mirror orientation is a very uncommon problem, but one which may occur—it did in our laboratory. During the preparations for a large study (310 participants in six days using four tower-mounted systems), we were cleaning the mirrors of the systems, and happened to put one of them upside down. Now, the mirror that is part of many eye trackers, is covered on one side by a thin layer that reflects infrared light but lets through visual light. It is usually difficult to place the mirror upside down, but when it happens, by accident, in the hurry of the last preparations, the infrared light is spread into four corneal reflections along the vertical axis, as in Figure 4.17(a). The diagnosis was difficult, but once the problem was found, the solution was easy.

Specks on mirrors may give rise to weak reflections in the eye image that reduce the contrast and increase noise levels, as shown in Figure 4.17(b). The solution is simple: clean the mirror.

The examples above should be ample proof of the importance of the user being able to see the eye video and influence its quality. Hiding the eye video from the user, and relying on automatic set-up of the eye video, is a current trend in manufacturer development. It is an understandable ambition to attempt to make this part of data recording more easy, but if this automatization is not properly done, it can make eye camera set-up and recordings more difficult, give data of a poorer quality, and leave the user with fewer clues on how to improve data quality.



(a) Four corneal reflections along the vertical axis, as a result of an infrared-reflecting mirror turned upside down.

(b) To the left of the pupil, a weak reflection is visible, equal in size to the pupil, that reduces the contrast and will increase noise levels in recordings. Clean the mirror from fingerprints and specks.

Fig. 4.17 Mirror reflections and lowered contrast due to a dirty mirror.

The entire eye camera set-up and positioning of the participant should take no more than a minute or two once you start to gather experience.

→ Βαθμ-ονομ-ηση

4.5 Calibration

Individual calibration of each participant is necessary for a variety of reasons. For instance, the eyeball radius varies by up to 10% between grown-ups, and it may also have different shapes. Glasses alter the size of the eye in the eye video image. These variations change the geometrical values underpinning the calculation of gaze direction (Hammoud, 2008).

If you are using a head-mounted system out of the laboratory, see to it that the level of ambient infrared light is as low and stable as possible. If you calibrate in an area with a lot of infrared light, and then record in a darker area, the contrast between pupil and iris may change and result in poorer data, unless the camera compensates and the system dynamically resets the contrast thresholds. For maximal data quality, always calibrate in the same luminance conditions as you will have during data record, and minimize luminance variation in the trials (see Chapter 2).

4.5.1 Points

Calibration is typically made on a 2D area which has a number of predefined calibration points: 2, 5, 9, 13, and 16 are common numbers. Calibration points should cover the area where the relevant stimuli will be presented, whether it is a monitor, a scene video from a head-mounted system, or even just a part of a screen. Using the pupil and corneal reflection positions recorded in calibration points with known coordinates, the recording software can calculate a function to estimate any given location on the stimulus, given the extracted pupil and corneal reflection positions. Systems with just one or even no calibration points are being developed. By calibrating the system on just a few calibration targets, the system can then fit a function that allows it to interpolate between all intermediate positions and also extrapolate to positions outside the calibration area. It should be noted that accuracy is better within or close to the calibration targets and the stimuli should preferably be appearing within the area encompassed by the calibration points.

Points should be small and visually salient, so that participants gaze at them as exactly as possible during calibration. Any misalignment of gaze towards the actual calibration point

pil, a weak reflection is
the pupil, that reduces
crease noise levels in
mirror from fingerprints

hould take no more than

f reasons. For instance,
may also have different
se variations change the
ammoud, 2008).
ee to it that the level of
e in an area with a lot of
upil and iris may change
stem dynamically resets
in the same luminance
nce variation in the trials

f predefined calibration
ts should cover the area
or, a scene video from a
pil and corneal reflection
e recording software can
given the extracted pupil
bration points are being
s, the system can then fit
ions and also extrapolate
uracy is better within or
ppearing within the area

gaze at them as exactly
e actual calibration point

---

will introduce a corresponding offset in your data. Some laboratories use high-frequency Gabor patches that are only visible when participants look straight at them. Animated calibration targets seem to work better than static ones, as participants gaze at them as long as they are animated rather than looking around for the next target. Having the participant mouse-click on very small points rather than just looking at them may increase accuracy even further. For small children, it can be an advantage to exchange the standard point for objects that are fun or somewhat familiar to them: a yellow sun, a blue star, etc.

The position of points differs substantially between systems that give data files and those that only give gaze-overlaid video output, such as the head-mounted systems. For data files, the calibration points should be put on the stimulus surface (i.e. most commonly the computer monitor on which the experimental stimulus is displayed). For gaze-overlaid data, the calibration points should be in the coordinate system of the scene video camera. When calibrating the scene video of a head-mounted system, a laser pointer is commonly used to project an actual target overlapping the calibration coordinates determined from the scene camera. Alternatively, it is possible to position the scene camera at an exact position where known physical targets will coincide with the calibration targets of the system.

4.5.2 Geometry

There are a number of geometry settings that the calibration routine needs access to: the size (and position) of the calibration screen in pixels and/or millimetres, the precise position of the calibration points, and the distance from the eye to the monitor. These values can be set manually with some eye-tracking systems, but are automatically estimated or measured in some of the remote systems.

Monitor distance affects data quality in corners, and what really matters is how much of the visual field is covered. Unless the participant has difficulties accommodating, there is little benefit in presenting on a 24 inch screen at a position chosen so that is covers a visual angle equivalent to a 17 inch monitor at a shorter distance. You have to find the proper compromise between the position when the corneal reflection is lost in top corners, when the pupil is covered with eyelashes at the bottom corners, and between the distance and height of the monitor in the overall set-up. Even with a very good set-up, some participants' eyes will nevertheless cause problems in bottom and top corners.

4.5.3 The calibration procedure

Always test the quality of your eye video before calibrating. Let the participant look at the corners of the calibration area and watch whether the eye video looks good with a stable pupil and corneal reflection in all corners. Then you can calibrate.

The participant looks through all calibration points, and at each point the eye-tracker software samples a few hundred milliseconds of data. The participant in Figure 4.18 looks at nine different points. Notice how the spatial relation between pupil and corneal reflection differs in the different gaze directions. It is of vital importance for the quality of your data that the full and correct pupil and corneal reflections are selected at every single calibration point. An eye video set-up with the images in Figure 4.18 would give a successful calibration and good quality data.

The progress through the points, as well as acknowledgement that the sampled data are valid, can be done *manually* either by the operator or by the participant, or *automatically* by the recording software. The current trend among manufacturers is to provide automatic calibration as their default. Tests that the authors have carried out with close to 60 participants in each of the three conditions (i.e. operator controlled, participant controlled, auto-

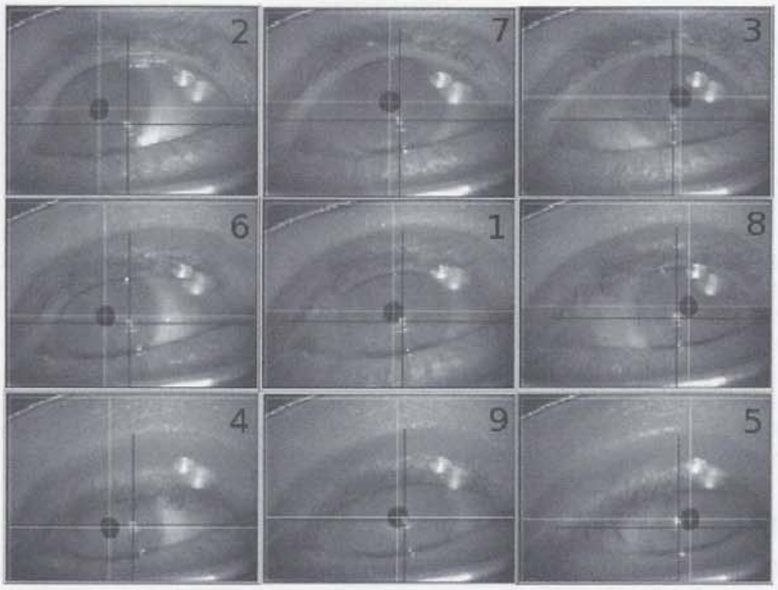


Fig. 4.18 Participant (with glasses) looking at the nine points of a calibration screen in the order given by the numbers 1 to 9. The bright reflections are always out of the way for successful detection of pupil and corneal reflections, although close in calibration point 3.

matic) showed that data quality (precision and accuracy) is superior for *participant-controlled* progress and acknowledgement of calibration points (Holmqvist, Nyström, & Andersson, 2011). In other words, participants know best themselves when they are looking at a point, and the researcher can hand over the calibration process with confidence.

Irrespective of calibration method, the participant must cooperate in looking at the calibration points, and they virtually always do. Exceptions may include children or animals who lose their interest in the calibration procedure after a few points. Patients who have difficulty keeping their eyes still, e.g. due to congenital nystagmus (common with albinos), are also difficult to calibrate.

You may sometimes miscalibrate participants simply because they happen to look in the wrong direction while you or the eye-tracker acknowledges one of the points. Sometimes, a participant tends to fixate the calibration points for a very short time, not knowing that you need a second or so to verify the image in the eye video and press the acknowledgement button. Other participants may make involuntary so-called square wave jerks during calibration (see Figure 5.30), giving an offset because the participant looked away slightly from the calibration point just at the moment that point was confirmed in the calibration procedure. To reduce the danger of such errors, always instruct your participants to keep their gaze fixed in the centre of each calibration point until it disappears, and watch out for participants from groups that are known to have a higher rate of square wave jerks (p. 407).

4.5.4 Corner point difficulties and solutions

Many systems allow you to watch the eye camera image while calibrating, so you can see that the eye is still, that it is looking in the right direction, and that corneal reflection and

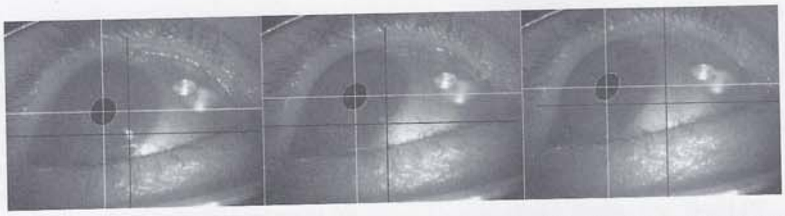


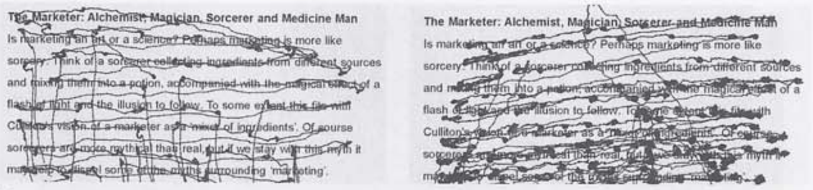
Fig. 4.19 More and more extreme gaze directions towards the upper left calibration corner. The first eye is OK, the second is dubious, and the third will give a considerable offset in data, as in Figure 4.20.

pupil are correctly identified, as in Figure 4.18. The most difficult calibration points are the corners. *Bottom corners* are problematic if your participant has downward eyelashes partially or completely covering the pupil, when looking down at the bottom corners. In that case, there are three things that you can do:

1. Ask the participant to open his/her eye. This will temporarily keep the eyelid or lashes from blocking the eye feature detection and lets you proceed with the calibration. But during recording, when your participant is fully occupied with the task and no longer thinks about holding the eye open, you will get optic artefacts, offsets, and data loss at that bottom corner. This may be a viable option if the problem only concerns the corners of the screen and the interesting parts of the stimuli are mainly located in the centre.
2. Acknowledge calibration of that point even if half the pupil is covered. This is gambling that your participant's pupil will also be half-covered during recording. Your data will have a smaller offset compared to the corners in case one, but the poorer precision and accuracy will remain, and will likely contain optic artefacts. If you are planning to do area of interest-based analysis this may be OK, but fixation and saccade measures will be affected.
3. Go back and set up the eye camera to film from further below, or ask the participant to curl the eyelashes. This takes more time, but it is the only guarantee of good data. With a remote, try tilting the camera or moving it closer or further from the participant to change the angle.

The *top corners* are usually easier, but they have a problem of their own, illustrated by the eye images in Figure 4.19 and the resulting data in Figure 4.20. When the corneal reflection moves across the border of the iris, it changes position and form. If you calibrate with a corneal reflection outside of the iris, then you will later have a considerable offset for data samples in that corner. There are of course large individual differences between participants as to the gaze direction at which this corneal reflection leaves the iris. For some participants with a narrow eye opening, the corneal reflection is instead covered by the lower eyelid. In any case, there are three things you can do to remedy this situation:

1. The quick solution is—if your recording software allows it—to move the calibration point further in, and ask your participant to look at the new position, hopefully thus moving the corneal reflection back into the iris. This solution will give precise and fairly accurate data when recording, especially if the task normally does not require the participants to look at extreme angles but rather in the centre of the screen.
2. You could also move your eye camera and/or infrared illumination so as to film the eye from further up. This takes more time, and you also run the risk of creating problems at the bottom corners if you change the camera angle too much.



(a) Upper right calibration point

(b) Upper left calibration point

Fig. 4.20 Offsets—poor accuracy—resulting from calibrating participants whose corneal reflection had left the iris, as in Figure 4.19. In 4.20(a), the upper right calibration corner was problematic. In 4.20(b), the upper left calibration corner. These errors are easy to spot in a highly structured design and task, such as in reading, whereas in other tasks it will be much more difficult to identify these offsets.

3. The most proper solution is to position the stimulus (monitor) so that both top and bottom corners are OK for the large majority of participants. This you have to do when building the recording environment and setting up the geometry of your experiment. A greater distance from the participant to the monitor will lower the maximum visual angles required and alleviate these problem, but it will also lower the difference in angles between different areas of interest and lower the accuracy and precision.

4.5.5 Calibration validation

It is important to never take a calibration at face value. A participant may have shifted his eyes just as you or the system captured the eye as representative for that particular calibration marker. The simplest solution is, right after the calibration, to show an image containing the calibration markers and ask the participant to look at the markers while you verify them visually on the control computer.

Some systems provide a numeric accuracy value of the average deviation between markers and gaze position. This accuracy value should be reported when you submit your article or report, but in the past this has been unusual. High-end systems exhibit values around 0.2°, whereas a maximum average deviation of 0.5° should be demanded for most studies. An average deviation of 1.0° would be unacceptable for instance for reading research investigating preview benefits and word landing positions, and some remotes produce 1.5° or larger average offset. If you use this number, be sure to relate the reported visual degree value to your particular stimuli display: how much off can the data be before your analysis suffers? Recalibrate until the validation values is below your required accuracy. Only then start the data recording.

It is important to keep in mind the difference between the estimated accuracy and the real accuracy. Even if you perform a validation, the validation points will also suffer from a random or systematic error. A realistic goal would be to have the offsets for the calibration/validation targets as small and random, i.e. offset equally in all directions, as practically possible. For studies that require very high accuracy, there exist correction methods that can be used immediately subsequent to calibration. For instance, Santini, Redner, Iovin, and Rucci (2007) describe a method in which raw data samples are shown on top of the calibration points, and a joystick is used to alter the underlying transformation matrix so that data are right above the points.

Although precision is just as important to validate as accuracy, in practice precision validation is very uncommon. An exception is Santini *et al.* (2007).

4.5.6 Binocular and head-tracking calibration

There are two special cases in which additional calibration work is necessary: binocular calibration, and calibration using head-tracking based eye-trackers.

Liversedge, White, *et al.* (2006) recommend a separate calibration for each eye when using (true) binocular recording and you are interested in investigating the disparity between the eyes. Calibrating both eyes at once may give an erroneous (absolute) disparity value, because the calibration routine believes that both eyes are directed towards the calibration point, when in fact one eye may be off. Other researchers, for instance Nuthmann and Kliegl (2009) argue that making separate calibrations of each eye is ecologically invalid, and report disparity measures even after simultaneous binocular calibration. Clearly, the relative disparity, and changes in disparity, will still be correctly measured.

When using head-tracking systems, additional calibration of the combined measurement system is typically necessary, but we refer to manufacturer manuals for details.

4.5.7 Calibration tricks with head-mounted systems

This section presents four calibration tricks that can be used on versatile head-mounted video-overlay eye-trackers to improve their performance. This concerns eye-trackers mounted on and moving with the participant, but otherwise working on the same eye-tracking principles described in this book.

Since calibration for such a system really only concerns calibration points in the scene video—the physical calibration points are only there to make the participant look at what corresponds to the right point in the video—it is possible for the participant to turn his/her head just as much as they want. This can be used in cramped places with partially unfortunate lighting conditions, such as sunlight, where there only is a single good physical target available. Choose one single physical calibration point so that the eye video looks good when the participant is looking in that direction. Then, for each calibration point, turn the participant's head so that the single physical calibration point is aligned with the calibration point in the video. If the participant continues to look at the physical point, he will in effect look in the correct directions for calibration, and the gaze-overlaid video data will be fine.

Another trick is to make a head-mounted system into an eye-tracker that yields a data file with absolute coordinates for an easier analysis. To do this, lock the head-mounted eye-tracker onto something, so it does not move relative to the stimulus monitor. You would probably have to build some construction. If successful, the eye video coordinates of your eye-tracker will be possible to use as coordinates in the stimulus monitor, and you can use the data file for fixation analysis, area of interest analysis, etc., which is currently not possible when your participant is moving around. There may be cases when such an approach is the easier option, e.g. in very specific environments when your standard remote or tower-mounted eye tracking will not fit, such as in a vehicle.

Yet another trick is to tilt the calibration area, if your system only supports a single stimulus area or plane. Suppose you are studying a car driver's gaze to interior controls on the dashboard, and also want to get high quality data for the outside traffic. If you calibrate all points outside the car, you will have a considerable parallax error (p. 60) for gazes inside the car. If you calibrate inside the car, you will have parallax error for outside targets. However, the calibration plane does not have to be orthogonal to the line of sight. Put the lower calibration points on the dashboard, and the upper points 10 metres in front of the car, to get a (more or less) parallax-free precision both on the dashboard and on the outside of the car.

A final tweak to get good data is when you have already calibrated, but notice that the gaze cursor is systematically offset in one direction or another due to movement of the head-band or the helmet, then it is possible to ask the participant to keep the eyes still at a particular position and then gently shift the scene camera or the entire system so that the offset disappears. This will allow you to fine-tune your calibration without re-doing the complete calibration, which may be an option in the field. Always do a validation test after this adjustment so you can determine later, during post-processing when you are looking at the recorded data, whether to include this data or not.

4.6 Instructions and start of recording

The last thing you should do before recording is to instruct your participants about the task. Be precise, and use exactly the same instructions for all participants. A written instruction is usually more similar between participants than a verbal instruction. If your participants have different conceptions of the task, they are very likely to behave differently. Task instruction has a very strong influence on eye-movement behaviour, as elegantly shown by Buswell (1935, p. 136ff) and Yarbus (1967, p. 174ff), and later largely replicated by Lipps and Pelz (2004) and DeAngelus and Pelz (2009). Note that when recording participant speech, additional specific instructions may be necessary.

If a computer program shows your stimuli, have that program start and stop your recording, as well as synchronizing it with changes in the stimulus display. Starting and stopping manually requires some amount of post-recording cutting of the data files. Synchronizing change of stimulus manually gives errors anywhere from 500 ms up to many minutes (if the operators are distracted). Sometimes manual start of recording is necessary, however, for instance, for field recordings in the real world.

Manipulating the eye camera set-up during recording can be necessary when the eye slowly disappears from the eye video, usually as a consequence of a head-mounted eye-tracker slipping or the participant slumping together slightly when relaxing. It will always disturb the participant, however, and often introduces measurement error. Some eye-tracking systems have built-in tools for quick recalibrating, which can be used in a pause between trials. On some older systems, recalibration was required often, such as once per line of read text. Newer systems with stable tracking can be left running for a long time without recalibrating or drift correction. However, if you know your system has drift issues, i.e. when the offset increases over time, then it is important to plan for frequent drift corrections between trials or blocks.

During recording, keep your eye on the eye video, and preferably keep a record of what happens: is the data good in the entire visual field? Is there a downward offset in the upper left corner? Is there a split corneal reflection problem? Is the data stable, or are there many optic artefacts? Taking such quality notes will be important for later interpretation of the data. Some of the problems can be remedied after recording, but some should be taken care of during recording. You can improve the eye video during recording: changing the focus of the camera and altering the threshold for pupil identification typically alters you data very little. But do not change the camera angle towards the eye, or the position of the infrared illumination. Such alterations will give you a permanent offset in your data, unless followed by immediate recalibration.

4.7 Auxiliary data: recording

This section covers the concurrent recording of auxiliary data along with eye movements and therefore may be regarded as optional for those solely concerned with learning the techniques of eye-movement recording. If you record only eye-movement data, you can skip to p. 139, but as we noted on pp. 95–108, there are many reasons why you may want to combine eye-movement data with other types of recordings. Here we will describe how to make combined recordings in three different types of technical set-ups:

Non-interfering set-ups *Reaction time tests* and *galvanic skin response* (GSR) are both technically simple and do not interfere with the eye-tracker in the sense that they do not cause reflections in the eye movements, large head movements, or other large sources of error.

Interfering set-ups *Motion tracking*, *EEG/ERP*, and *fMRI* are technically and methodologically challenging, and may interfere with the eye-tracker (or the eye-tracker may interfere with the auxiliary hardware). For example, a contact-based eye-tracker will generate noise if the EEG electrodes touch the eye-tracker or if the eye-tracker generates much electromagnetic noise.

Verbal elicitations *Verbal data* may interfere with the eye-tracker, or vice versa, and may require additional instructions in order to acquire a high enough data quality (pp. 105–106), but is such a common auxiliary data source that it warrants its own section.

4.7.1 Non-interfering set-ups

Reaction time tests are often measured using a so-called *button-box* with a limited number of buttons, and giving the participant the instruction to click the correct button—right for yes, and left for no, for instance—as soon as they know the answer to the question. Using a third-party stimulus and synchronization software, reaction time latency measurements are easy to combine with eye tracking. Unless the click causes vibrations in the eye-tracker (p. 35), there is no interference between reaction time tests and eye tracking. Some manufacturers (e.g. SR Research) provide software for reactions time tests to be incorporated into the design of an experiment simply using the existing keyboard, but note that button boxes have better temporal characteristics, and other software specifically designed for the kinds of reaction time tests used in studies on attention can easily be used in conjunction with eye tracking.

Galvanic skin response (GSR) measures the electrical conductivity of the skin using electrodes which are usually put on one or two fingers of the participant. The variation in GSR signal corresponds to the autonomic nerve response as a parameter of the sweat gland function. Most GSR systems can be run at the same sampling frequency as eye-trackers, but stimulus presentation and synchronization may have to be dealt with using third-party software. There is no mechanical or electromagnetic interference between the GSR and eye-tracking systems.

4.7.2 Interfering set-ups

Eye-trackers that work for participants inside the intense magnetic field of the fMRI tunnel have been sold for quite some time, but the market is small, and the benefit of eye tracking to fMRI (or vice versa) is still unclear to many users.

An eye-tracker that works in a strong magnetic field must be built differently from the eye-trackers we saw in Chapter 2. All electronics must be at a far enough distance from the magnet, so fMRI-compatible eye-trackers use long-range cameras and mirrors close to the participant, as in Figure 4.21.

Initially, eye-movement recording in combination with *EEG* was done using an EOG (electro-oculography) system. EEG sensors and EOG sensors are the same, only placed differently. As eye movements and blinks give rise to problematic artefacts in EEG data, many current-day EEG systems include EOG sensors. The EOG data can then be used to filter the

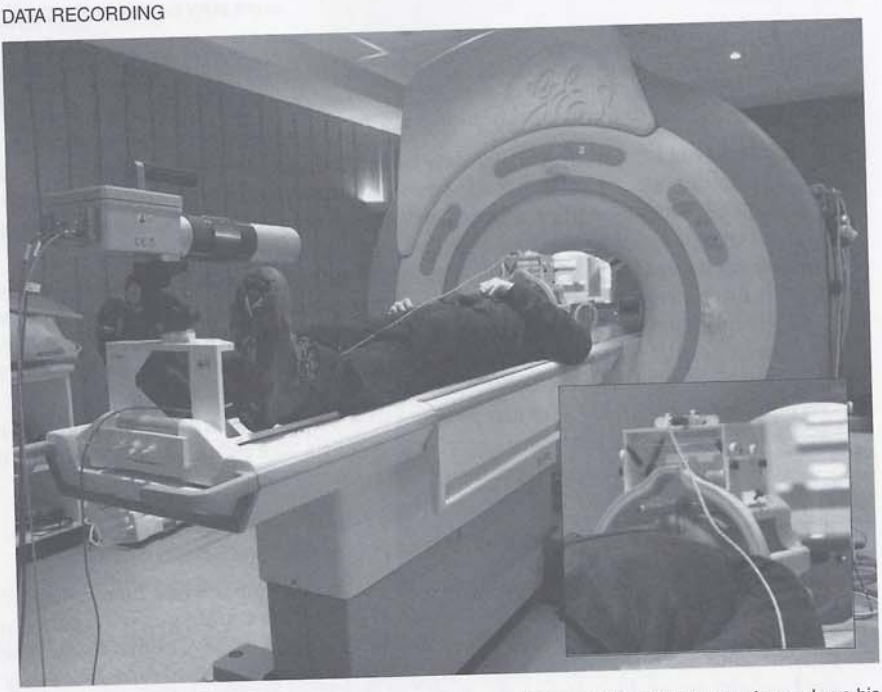


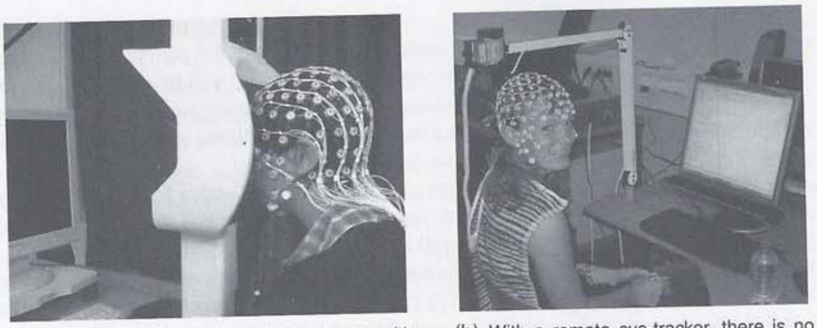
Fig. 4.21 Long-range eye-tracker for use in fMRI environment. The participant has two mirrors above his face: one mirror for the eye camera to film through, the other for projecting the stimulus image. Image used with kind permission from SMI GmbH.

eye movements and blink signals in the EEG data (Jung *et al.*, 2000). Even now, EOG data is used to remove the movement and blink artefacts from the EEG signal, using what is known as the *surrogate multiple source eye correction* (Berg & Scherg, 1994). Here, a number of reference eye movements are recorded before the actual experimental data. a PCA (principle component analysis) reveals the artefact topographies, and the waveforms from these can be modelled in the experimental data so that overcorrection is avoided. After subtraction of the artefact waveforms, the EEG data will be cleaned from oculomotoric artefacts.

In later years, EEG recordings were also combined with video-based eye-movement recordings. Figure 4.22 shows how a tower-mounted high-precision eye-tracker and a remote eye-tracker are used in combination with EEG. In neither of the two studies where these combinations were used could we detect electromagnetic interference between the two systems, not even in the tower-mounted eye-tracker where the electronics are really close. We did see some mechanical interference in the tower-mounted eye-tracker, however (Simola, Holmqvist, & Lindgren, 2009). Leaning a forehead covered with EEG sensors against the head stabilizer of the tower-mounted eye-tracker was easy for some participants, but it was difficult for some to keep still. The eye-tracker tended to lose data as participants changed position and the eye moved out of camera view, but EEG data appeared not to be affected, either from eye-tracking equipment or TFT monitors. CRT monitors may have to be shielded.

The EEG system should be mounted and calibrated on the participant first, as it will always take longer than the eye-tracker set-up and calibration.

Motion trackers exist both as magnetic and optic varieties, but optic ones are more common and versatile. Since they also operate with infrared light and reflections, optic motion



(a) When the EEG net is in close contact with the head-stabilizing part of a tower-mounted eye-tracker, data in both systems are sensitive to slippage: the eye may move outside of the camera image as the participant attempts to find a comfortable position. The foremost sensors on the cap or net may slip on the forehead. For participants who can remain still, data is excellent, however. The 128-channel HydroCel EGI Geodesic SensorNet and the SMI 240 Hz HiSpeed eye-tracker.

(b) With a remote eye-tracker, there is no contact between the two systems. The participant is free to move her head, and there is less physical interference. Remotes invariably give a poorer data quality than tower-mounted systems, however. The 128-channel HydroCel EGI Geodesic SensorNet and the SMI RED IV 50 Hz remote eye-tracker.

Fig. 4.22 EEG and tower-mounted versus remote eye tracking.

trackers could potentially disturb the eye-tracker, or in odd circumstances the other way around, but in our experience this rarely happens. Directing the motion tracker cameras differently and turning down the light intensity of their lamps may remedy interference. Many motion trackers run at the same range of sampling frequencies as eye-trackers.

4.7.3 Verbal data

Verbal data can be recorded *concurrently* during eye-tracking recording or *retrospectively*, after eye-tracking recording (pp. 101–105). The point in time of recording the verbal data may influence the quality of the eye-tracking data dependent on the *type of eye-tracker*, as we will see in the following. Thus, before deciding when to record the verbal data, you have to consider which type of eye-tracker you have and what kind of cutbacks you are willing to accept in the quality of your eye-tracking data. If you decide to record the verbal data retrospectively, the verbal data should be cued by a replay of the participant's eye-tracking data to stimulate the participant's memory of their thoughts (compare Van Gog, Paas, & Van Merriënboer, 2005). Consequently, you have to consider *how to present* these eye-tracking data to the participant. Irrespective of when you record the verbal data, the crucial issue is how you *instruct* your participants to provide unbiased verbalizations (p. 105).

Task interference and tracking robustness

When recording verbalizations *concurrently*, make sure that the participant is able to speak freely. Remote, tower-mounted, and head-mounted eye-trackers interfere differently with the speech movements of the participant, as will be stressed in this section. Having participants talk while being eye-tracked also introduces a somewhat poorer quality in the eye-movement data, in the form of increased noise levels, optic artefacts, or inaccuracies, simply because participants move more when they speak, and the various head movement compensations can not always keep up.

When recording with a *remote* eye-tracking system (without adding a forehead rest), the influence of speech movements on eye-tracking data quality will be moderate. For instance, fast head movements may not be compensated for, and the physiology around the eye may change as speech muscles move. The freedom of the participant to move their head is likely to give better concurrent speech quality, however.

When using a *head-mounted* system (or a remote system with a forehead rest), concurrent verbalizations do not disturb the eye movement recordings as much. In this case, an additional microphone needs to be used, which will be attached to the participant.

A *tower-mounted* high-end system does restrict head movements, but care must be taken not to restrict jaw movements during concurrent speech; such precautions may however jeopardize eye-movement data quality instead. Retrospective verbalizations will provide the best eye-tracking data.

When recording *retrospective* verbalizations, the major technological challenge is to minimize the delay from the end of recording eye-movement data until recording of retrospective verbalizations is started. A longer delay means poorer quality of verbal data, as the participant's memory quickly deteriorates. The bottleneck is the possibility of showing the eye-movement data very quickly to the participant, and start sound recording fast. Different manufacturers offer varying support for this. Some software allows for a direct and quick replay of recorded eye movements. Others require the researcher to export the recorded eye-movement data from the recording software and then import these data to the analysis software before they can be shown. Another important technological feature for retrospective verbalizations is the ability to replay the recordings full-screen, so that the participant is not distracted by the user interface of an eye-tracking data analysis software. Further potentially important technological features include being able to display the participant's eye movements with transparent dots and scanpaths, so that important parts of the display are not occluded.

Note that participants may not only speak about their own eye-movement data, but may very well point to them, using deictic references like "there" and "it" towards the indicated item. With frequent pointing comes a need to record video along with sound, and to transcribe pointing activities.

How to display the eye-movement data when recording cued retrospective verbalizations

When recording verbal data retrospectively, the aim is to elicit thoughts as close as possible to the actual thoughts during the experiment. This leads to better results—in terms of less forgetting and less fabricating—when participants' verbalizations are cued (Van Someren *et al.*, 1994).

Participants' eye-movement data may be shown to him either as *static* visualizations (showing a longer sequence at once), or *dynamic* visualizations (momentous). It is likely—although not proven—that these two types of eye movement visualizations elicit different verbal data from the participant, since they deliver different types of information. We would argue that each visualization has its own advantages depending on the task participants have to accomplish during the experiment and your research question, but academic research into this method is lacking.

Static visualizations of eye movements give an overview over a longer span of time, typically the entire trial (webpage, learning instruction, or sentence to be translated). The temporal order of a short scanpath can be shown by order numbers of the fixations. For longer sequences, the temporal extension during which the scanpath is visible may be interactively altered to reveal the order. Having an overview can be expected to prime the participant to comment on relations between inspecting different areas, like why there are a lot of saccades between two areas, but not between others, or explaining the inspection of areas with lots of fixations. However, it is very likely that the order of inspection will be neglected, and that only areas that were intensively inspected are mentioned.

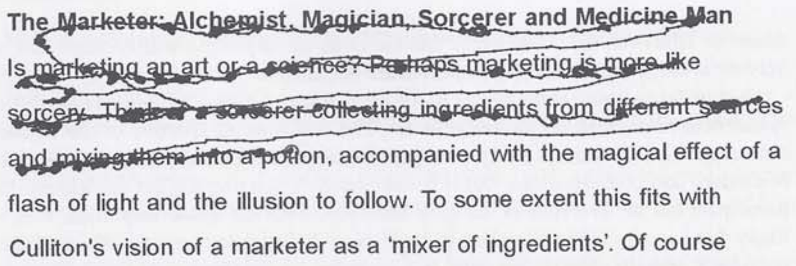
A dynamic visualization of eye movements instead shows the momentous dynamics of visual behaviour, with no overview at all. The real-time perspective of the dynamic eye-movement visualization is likely to trigger participant comments on the instant decisions, hesitation, and considerations, but it is also more demanding for the participant. Since the participant has no overview of his gaze behaviour over the whole inspection, it is very unlikely that he will comment on larger attention shifts in space over time. Since dynamic eye-movement visualization is also quite fast, reducing the playback speed may be necessary to allow the participant time to reflect on what he had been thinking during the task (Kammerer, Bråten, Gerjets, & Strømsø, 2010). However, slowing down the recording is not always appropriate. For stimuli that depict dynamic content a slowed speed would change the features of the stimulus itself. Another possibility is to reiterate the dynamic eye-movement visualization several times to give the participant several chances (Jarodzka, Scheiter, *et al.*, 2010). This possibility, however, is difficult in terms of analysis later on. For instance, what if a participant, who viewed the eye-movement visualization several times, makes a more detailed description of the stimulus than another participant who viewed his eye-movement visualization only once? In such cases, at least time on task has to be added as covariate into the analyses. A third possibility is to enable a self-paced, participant-governed eye movement visualization; this might help the participant even more to deal with the complexity of the presented data (Van Gog, Paas, Van Merriënboer, & Witte, 2005). Again, very different durations of inspecting eye-movement visualizations may however occur. Moreover, it is likely that not all participants are equally able to estimate their need to adapt the eye-movement visualization. This in turn would result in diverse quality of verbal data. A final concern would be if dynamic visualizations put social pressure on the participant to produce more, as more events unfold in the visualization, yet the participant has nothing to contribute. Then this social pressure may result in a participant confabulating to escape the pressure.

How to elicit verbalizations from participants

When recording verbal data as thinking aloud the precise instruction to the participants is of great importance (p. 105). The instruction for the participant is done in three steps: *instructing*, *training*, and *reminding*. It is important that the participant feels comfortable during recording. Training before the actual recording helps the participant become familiar with the situation and task. Moreover, it helps when as few people as possible are in the recording room, so the participant does not feel monitored. In case the participant stops verbalizing his thoughts, the experimenter has to gently *remind* him to go on.

4.8 Debriefing

After all data collection is done, debrief the participant. It may be a good idea to start by asking the participant if he can guess what the study was about. You might do this to see if the actual purpose of the study is (too) transparent. Then, tell the participant what the study was really all about. This is particularly important if the experiment was based on deception or the nature of the experiment is likely to influence the participant's belief in his abilities, such as a reading test (see e.g., Sharpe & Faye, 2009 for a discussion about debriefing). As part of debriefing, you should also ask them to sign your written informed consent document, unless this was done before the data recording. Inform the participant that his recorded data will be anonymized and that he can withdraw data at any point. Make sure to underline that the participant is welcome to get in touch anytime, with any questions he may have. He might



The Marketer: Alchemist, Magician, Sorcerer and Medicine Man

Is marketing an art or a science? Perhaps marketing is more like sorcery. Think of a sorcerer collecting ingredients from different sources and mixing them into a potion, accompanied with the magical effect of a flash of light and the illusion to follow. To some extent this fits with Culliton's vision of a marketer as a 'mixer of ingredients'. Of course

Fig. 4.23 Good data quality. Raw samples (1250 Hz), showing fixations as blobs and saccades as lines.

not want his data to be used in the overall analysis of the study, and if this is the case ethical principles dictate that this request must be respected. Also during debriefing, see to it that your participant has your contact information. In practice, previous participants almost never get in contact with us, but they should be able to if they wish.

If possible, procure an inconvenience allowance for participants, but beware of tax regulations and complex administration. Some tax legislations force experimenters to formally hire their participants above a certain reward level. This is important to check before you start giving out money. Finally, ask the participants not to reveal the purpose of the study, should they bump into someone they know, who might be your participant the next day, and who should not know about its purpose.

If you have done everything right, your participant will leave feeling at ease, happy with having contributed to research.

4.9 Preparations for data analysis

After the data recording is finished, backup and classification of data proceeds. We now have a set of files containing data that will be pre-processed and then analysed.

4.9.1 Data quality

As we pointed out on page 87, the first thing to do after data are backed up is to investigate data quality. Data quality can to some extent be measured by recording participant fixations on a collection of points just before, during, and after recording, and then calculating precision and accuracy values in each point. Usually, however, data quality is estimated by visual inspection of scanpath visualizations, of which many have been shown in this chapter, and velocity over time graphs, which we introduced on page 48, and which will be much used in the next chapter.

When exhibited as a scanpath, good data quality looks as in Figure 4.23. When we recently recorded reading data from 310 participants—of which one is shown in Figure 4.23—around 90% of the data have this quality. Fixations are more or less right on the word (so we can make area of interest analysis), and the precision is good enough for the fixation and saccade algorithms. The figure shows raw samples, because disturbances are more clearly seen—event detection smoothens scanpaths considerably, as we will soon see.

In some portions of your data, quality may be lower. A participant can have good data for some trials or parts of trials, and horrible data elsewhere. Data quality can also vary across the stimulus, being excellent in one part and awful in another. Whether the problem is an offset, low precision, optic artefacts, or complete data loss in one portion of the field of view, it matters, because the nature of your data quality determines its suitability for different subsequent analyses. For instance, the corner offset in Figure 4.20 does not prevent you from making a calculation of fixation durations, but area-of-interest order data based on word-position regions would be invalid. Therefore, it is useful to make data quality notes during recordings when you can see in real time what is going on— and use those notes to sort your data files to different analyses. Later, confirm uncertain judgments by inspecting raw scanpath and velocity-over-time plots. Some data files you will simply have to throw away, for some you can possibly do partial analysis—like the top half of the monitor when you had a drooping eyelid participant; maybe you can only calculate fixation durations for some recordings. Just remember that if you use data with poor quality for a dependent variable that it is not suited for, you will end up presenting false results in your papers.

How much of the data can you expect to have to throw away? In our experience, around 2–5% of the data from a population of average non-pre-screened Europeans needs to be dismissed due to participant-specific tracking difficulties. However, this number varies significantly: Schnipke and Todd (2000), Mullin, Anderson, Smallwood, Jackson, and Katsavras (2001) and Pernice and Nielsen (2009) report data losses of 20–60% of participants/trials, and Burmester and Mast (2010) excluded 12 out of 32 participants due to calibration (7) and tracking issues (5). The amount of data loss can be related to several factors such as the experience of the experimenter, the quality and flexibility (possibility of changing set-up of camera and mirrors) of the hardware, the complexity of the experiment, the calibration procedure, and how well the recording environment is controlled. Also, when combining EEG nets and tower-mounted systems, the data loss may increase to almost 1/3 of the data (Simola, Holmqvist, & Lindgren, 2009). Using an eye-tracker outdoors, near sunlight, and in cars may lead to a considerable data loss due to ambient lighting conditions. Wang *et al.* (2010) lost data from 32% of the participants in field car driving, for instance. However, these data losses should also be seen in a wider perspective, for at least two reasons. Screening participants beforehand so only those with no glasses, no lenses, wide-open eyes, not too bright iris, and upward lashes, are invited as participants, can give a 100% tracking robustness, but it should also be noted that this may involve recruiting a subset of the population that may not be representative. Additionally, reading about data loss in an article paper must not necessarily mean there were technological difficulties, but that the researchers were very picky about their quality criteria and rather removed data than include overly noisy or invalid data.

There is no standard criterion for *how* to select the data to discard, however. Criteria used in literature include, for instance, the percentage of zero values in the raw data samples, a high offset (poor accuracy) value during validation, a high number of events with a velocity above 800°/s, and an average fixation velocity above 15°/s (indicative of low precision). For an example of accuracy and data loss criteria, see Komogortsev, Gobert, Jayarathna, Koh, and Gowda (2010).

4.9.2 Analysis software for eye-tracking data

Analysis softwares naturally lag behind the research front. The ones that exist sometimes leave the researcher no other option than to program the analysis herself. Historically, analysis software have moved through three levels of sophistication.

- The simplest (and oldest) ones can import one data file (that is one participant) at a time, and include a calculation of fixations (using some algorithm), the possibility of drawing areas of interest on stimulus images, and a few visualizations—typically a scanpath or a proportion of gaze to different areas of interest—but only for one participant at a time. Randomization of stimulus images is not resolved. Up until around 2000, this is what most researchers had to work with, unless they programmed analysis themselves.

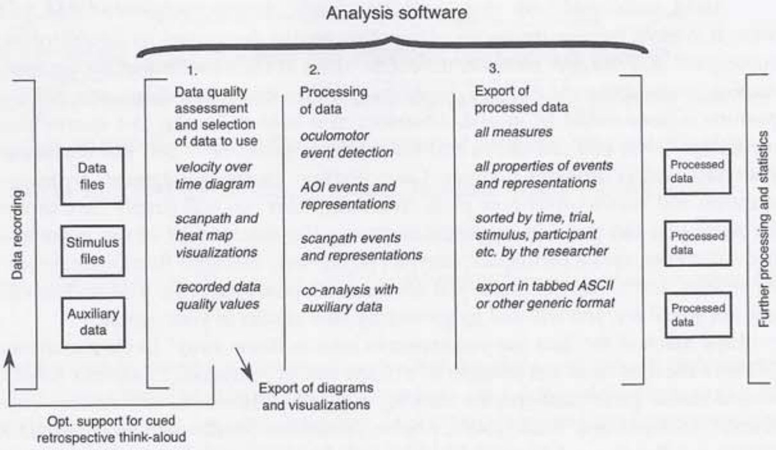


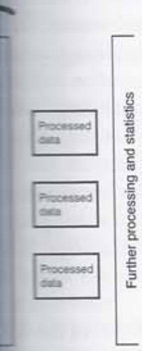
Fig. 4.24 An analysis software takes data files resulting from recording for an experiment. The three stages 1–3 correspond to a typical workflow in research, with italicized terms for visualizations, tools, and export possibilities. After data analysis, further processing and statistics take part in other, dedicated software.

- The middle level software is more integrated with recording and stimulus presentation programs. It can import a large number of stimulus images that have been used in an experiment, and resolve randomization. It exports a few (4–30) types of tables with data for several participants or several images at once. It is still more oriented towards visualizations than to export of structured data, however. These softwares first appeared in the early 2000s.
- The highest level of software is tightly integrated with recording and stimulus presentation software. The software allows for analysis of data from varying technical set-ups. It allows the user not only a choice between several fixation algorithms, but also powerful area of interest, scanpath, and attention map analyses. It allows the user to group participants, stimuli, and areas of interest according to the experimental design, supports the workflow suggested in this book, including the formatted export of a vast number (50–200) of research-relevant dependent variables into formats that can easily be processed further in a statistics package such as SPSS.

Figure 4.24 shows what three components you should expect in a commercial analysis software. First, tools for assessing the data quality and selecting which data has good enough quality to use. Second, several eye-tracking-specific processing tools, corresponding to Chapter 5–9, should be represented. Third, the export of structured data from the processing tools to statistical software is of great importance.

When we have interviewed them, the developers of analysis software have mentioned two different strategies. Some only provide the most basic output in stage 3 of Figure 4.24, arguing that researchers will be so inventive that new measures will appear all the time and seldom be reused, and that each user should implement any further analyses themselves. Others try to define and implement a set of often recurring analyses that many users can benefit from, which is a much larger undertaking.

This book covers the analysis options present in the most advanced types of existing analysis software. Many established eye-tracking laboratories do program their own analyses, however, either because the current software solution for their system is not ideal, not cutting



xperiment. The three
r visualizations, tools,
art in other, dedicated

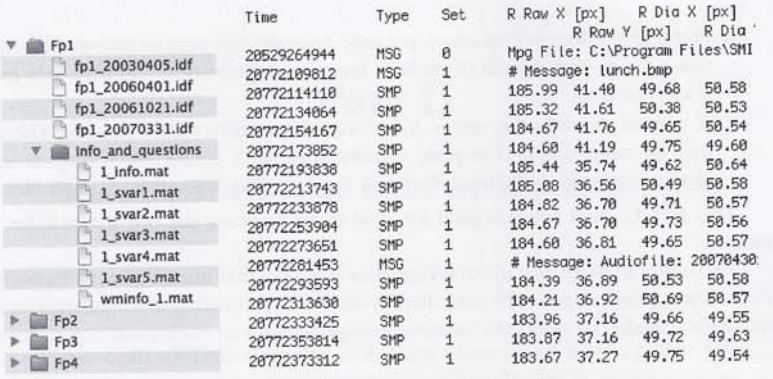
imulus presentation
ave been used in an
types of tables with
ore oriented towards
wares first appeared

d stimulus presenta-
g technical set-ups.
ithms, but also pow-
ws the user to group
imental design, sup-
ted export of a vast
rmats that can easily

commercial analysis
ata has good enough
responding to Chap-
the processing tools

have mentioned two
Figure 4.24, arguing
the time and seldom
emselves. Others try
ers can benefit from,

ced types of existing
n their own analyses,
not ideal, not cutting



```
▼ Fp1
    fp1_20030405.idf
    fp1_20060401.idf
    fp1_20061021.idf
    fp1_20070331.idf
  ▼ Info_and_questions
      1_info.mat
      1_svar1.mat
      1_svar2.mat
      1_svar3.mat
      1_svar4.mat
      1_svar5.mat
      wminfo_1.mat
▶ Fp2
▶ Fp3
▶ Fp4
```

(a) Data files.

```
Time          Type   Set   R Raw X [px]   R Dia X [px]
                                  R Raw Y [px]   R Dia
20529264944   MSG    0     Mpg File: C:\Program Files\SMI
20772109812   MSG    1     # Message: lunch.bmp
20772114110   SMP    1     185.99  41.40   49.68   50.58
20772134064   SMP    1     185.32  41.61   50.38   50.53
20772154167   SMP    1     184.67  41.76   49.65   50.54
20772173852   SMP    1     184.60  41.19   49.75   49.60
20772193838   SMP    1     185.44  35.74   49.62   50.64
20772213743   SMP    1     185.08  36.56   50.49   50.58
20772233878   SMP    1     184.82  36.70   49.71   50.57
20772253904   SMP    1     184.67  36.70   49.73   50.56
20772273651   SMP    1     184.60  36.81   49.65   50.57
20772281453   MSG    1     # Message: Audiofile: 20070430
20772293593   SMP    1     184.39  36.89   50.53   50.58
20772313630   SMP    1     184.21  36.92   50.69   50.57
20772333425   SMP    1     183.96  37.16   49.66   49.55
20772353814   SMP    1     183.87  37.16   49.72   49.63
20772373312   SMP    1     183.67  37.27   49.75   49.54
```

(b) Inside each eye-movement data file.

Fig. 4.25 Typical data after data recording is completed, but before analysis is started. Data often reflect the experimental design. In this case (a) shows four binary eye-movement files (`*.idf`) for each participant because there were four conditions, and several auxiliary Matlab (`*.mat`) files with data from questionnaires taken during the trials. Correspondingly, (b) shows that inside each eye-movement file, we find the data for raw samples, one for each line, time stamps, x and y coordinates in the stimulus for the right eye, horizontal and vertical pupil dilation, and on some lines, the onset of stimuli (`lunch.bmp`) and audio recordings.

edge, or they do not trust its output. Also, the particular analysis that the researcher wants may not be supported by the software. MatLab seems to be the dominating choice of programming environment for eye-tracking labs, but the freely available programming language Python and the open-source statistical language R are quickly increasing in number of users. Code can easily be shared among laboratories, and some code is made publicly available on the Internet. One absolute requirement for continuing with custom-made analyses is the ability to export into neutral file formats accessible by all software, such as tabbed ascii (.tsv). These files, with eye-tracking data, will look something like Figure 4.25.

When co-analysing eye movements with other data (p. 286), researchers typically use one analysis tool for each type of data, and then do an additional joint analysis in MatLab, R, or other software packages.

4.10 Summary

This chapter has discussed all the potential pitfalls you may encounter during data recording. Additionally, you should now know enough to be able to troubleshoot accuracy and precision problems, and be able to handle participants with more challenging eyes.

Some key insights include:

- Selecting the appropriate stimulus presentation solution for your experiment, knowing the benefits and drawbacks of each.
- Several degrees of freedom are available in most eye-trackers to reposition the system or angle it differently to overcome mascara problems and reflections in glasses.
- Simple tricks can be very effective, such as slightly defocusing the camera if the participant wears contact lenses, or simply having a make-up removal kit in the laboratory to be able to ask participants to remove mascara.

- Participant ethics is important, not only because you have to follow the rules, but because happy participants come back for more experiments, and may help recruit their friends as well.
- If needed, there is the option to record with auxiliary data sources, although these may be more or less disrupting or interfering with your eye-tracking set-up. EEG is challenging to include, but button-box reaction time tests are very unproblematic.

This and the plan you designed for your experiment are the raw material for your data analysis.

Recording high-quality eye-tracking data is a skill that must be learnt through training. Poor data quality can never be remedied by later data processing and statistical analysis, just like a good quality of your data can never compensate for an inadequate experimental design.