Why you need to care about design

1

1.1 Why experiments need to be designed

When life sciences students see the phrase 'experimental design', it can either ease them gently into a deep sleep or cause them to run away screaming. For many, experimental design conjures up unhappy memories of mathematics or statistics lessons, and is generally thought of as something difficult that should be left to statisticians. Wrong on both counts! Designing simple but good experiments doesn't require difficult maths. Instead, experimental design is more about common sense, biological insight and careful planning. Having said that, it does require a certain type of common sense, and there are some basic rules. In this book, we hope to steer you towards thinking more effectively about designing experiments.

Designing effective experiments requires thinking about biology more than about mathematical calculations.

So why are many life scientists so averse to thinking about design? Part of the reason is probably that it is easy to think that time spent designing experiments would be better spent actually doing experiments. After all, the argument goes, we are biologists so let's concentrate on the biology and leave the statisticians to worry about the design and analysis. This attitude has given rise to a number of myths that you can hear from the greenest student or the dustiest professor.

Myth 1 ***It does not matter how you collect your data, there will always be a statistical 'fix' that will allow you to analyse them.***

It would be wonderful if this was true, but it is not. There are a large number of statistical tests out there, and this can lead to the false impression that there must be one for every situation. However, all statistical tests make assumptions about your data that must be met before the test can be meaningfully applied. Some of these assumptions are very specific to the particular test. If you cannot meet these, there may be a substitute test that assumes different characteristics of your data. But you may find that this alternative test only allows you to use your data to answer a different scientific question to the one that you originally asked. This alternative question will almost certainly be less interesting than your original question (otherwise why weren't you asking it in the first place?). It's also generally true that well designed experiments require simpler statistical methods to analyse them than less well designed experiments. Further, there are some basic assumptions that apply to all statistical tests, and you ignore these at your peril. For instance, statistical tests generally assume that your data consist of what statisticians refer to as **independent data points** (more on this in Chapter 3). If your data don't meet this criterion then there is nothing that statistics can do to help you, and your data are useless. Careful design will allow you to avoid this fate.

Independent data points come from unconnected individuals. If the measured value from one individual gives no clue as to which of the possible values the measurement of another individual will produce, then the two measurements are independent.

Careful experimental design at the outset can save a lot of sweat and tears when it comes to analysing your data.

The group of experimental subjects that we use in our experiment (called **the sample**) needs to be representative of the wider set of individuals in which we are interested (called **the population**). One key way to achieve this is to ensure that each individual selected is not linked in some way to another individual in the sample, i.e. they are independent. For example, if we were surveying human food preferences, then gathering data from five members of the same family would not produce five independent data points. We would expect that members of the same family are more likely to share food preferences than two unconnected individuals, since family members often have a long history of eating together. Similarly, gathering data from the same person on five separate occasions certainly does not provide five independent data points, since a person's preference on one occasion is likely to be a good guide to their preferences a little while later.

Q 1.1 If we wanted to measure the prevalences of left-handedness and religious practices among prison inmates, what population would we sample from?

Q 1.2 If we find that two people in our sample have been sharing a prison cell for the last 12 months, will they be independent samples?

Myth 2: *If you collect lots of data something interesting will come out, and you'll be able to detect even very subtle effects.*

It is always reassuring to have a notebook full of data. If nothing else, it will convince your supervisor that you have been working hard. However, quantity of data is really no substitute for quality. A small quantity of carefully collected data, which can be easily analysed with powerful statistics, has a good chance of detecting interesting biological effects. In contrast, no matter how much data you have collected, if it is of poor quality, it will be unlikely to shed much light on anything. More painfully, it will probably have taken far longer and more resources to collect than a smaller sample of good data.

In science, hard work is never a substitute for clear thinking.

1.2 The costs of poor design

1.2.1 Time and money

Any experiment that is not designed in an effective fashion will at best provide limited returns on the effort and resources invested, and at worst will provide no returns at all. It is obvious that if you are unable to find a way to analyse your data, or the data that you have collected do not enable you to answer your question, then you have wasted your time and also any materials. However, even if you don't make mistakes of this magnitude, there are other ways that a poorly designed experiment might be less efficient. It is a common mistake to assume that an experiment should be as big as possible, but if you collect more data than you actually need to address your question effectively, you waste time and money. At the other extreme, if your experiment requires the use of expensive consumables, or is extremely time consuming, there is a temptation to make it as small as possible. However, if your experiment is so small that it has no chance at all of detecting the effects that you are interested in, you have saved neither time nor money, and you will probably have to repeat the experiment, this time doing it properly. Problems like these can be avoided with a bit of careful thought, and in later chapters we will discuss ways of doing so.

Similarly, it is not uncommon for people to collect as many different measurements on their samples as possible without really thinking about why they are doing so. At best this may mean that you spend a great deal of time collecting things that you have no use for, and at worst may mean that you don't collect the information that is critical for answering your question, or that you do not give sufficient time or concentration to collecting the really important information.

Don't be over-ambitious: better that you get a clear answer to one question than a guess at the answers to three questions.

Thus, while it is always tempting to jump into an experiment as quickly as possible, time spent planning and designing an experiment at the outset will save time and money (not to mention possible embarrassment) in the long run.

In science, as in life: more haste; less speed.

1.2.2 **Ethical issues**

If the only issue at stake in ill-conceived experiments was wasted effort and resources that would be bad enough. However, life science experiments have the additional complication that they will often involve the use of animals. Experimental procedures are likely to be stressful to animals; even keeping them in the laboratory or observing them in the wild may be enough to stress them. Thus, it is our duty to make sure that our experiments are designed as carefully as possible, so that we cause the absolute minimum of stress and suffering to any animals involved. Achieving this will often mean using as few animals as possible, but again we need to be sure that our experiment is large enough to have some chance of producing a meaningful result.

There are often many different ways that a particular experiment could be carried out. A common issue is whether we apply several treatments to the same individuals, or apply different treatments to different individuals. In the former case we could probably use fewer individuals, but they would need to be kept for longer and handled more often. Is this better or worse than using more animals but keeping them for less time and handling them less often? The pros and cons should be weighed up before the experiment is done to ensure that suffering is minimized. Issues such as this will be explored in more detail in Chapter 4.

Ethical concerns do not only apply to scientists conducting experiments on animals in a laboratory. Field experiments too can have a detrimental effect on organisms in the environment that the human experimenters are intruding on. There is no reason to expect that such effect would even be confined to the organism under study. For example, a scientist sampling lichens from hilltops can disturb nesting birds or carry a pathogen from one site to another on their collection tools.

We cannot emphasize too strongly that, while wasting time and energy on badly designed experiments is foolish, causing more human or animal suffering or more disturbance to an ecosystem than is absolutely necessary is inexcusable.

1.3 The relationship between experimental design and statistics

It might come as a surprise that we are not going to talk about statistical analysis of data in any detail in this book. Does this mean that we think that statistical tests are not important to experimental design and that we don't need to think about them? Absolutely not! Ultimately, experimental design and statistics are intimately linked, and it is essential that you think about the statistics that you will use to analyse your data before you collect them. As we have already said, every statistical test will have slightly different assumptions about the sort of data that it requires or the sort of hypothesis that it can test, so it is essential to be sure that the data that you are collecting can be analysed by a test that will examine the hypothesis that you are interested in. The only way to be sure about this is to decide in advance how you will analyse your data when you have collected it. Thus, whilst we will not dwell on statistics in detail, we will try to highlight throughout the book the points in your design where thinking about statistics is most critical.

There are two reasons why we will not concentrate on statistics in this book. The first is simply that there are already some very good statistics books available, that you can turn to when you are at the stage of deciding what test you are planning to do (see the Bibliography for a guide to some that we have used). However, the second, more important, reason is that we believe strongly that experimental design is about far more than the statistical tests that you use. This is a point that can often be lost among the details and intricacies of statistics. Designing experiments is as much about learning to think scientifically as it is about the mechanics of the statistics that we use to analyse the data once we have them. It is about having confidence in your data, and knowing that you are measuring what you think you are measuring. It is about knowing what can be concluded from a particular type of experiment and what cannot.

Experimental design is about the biology of the system, and that is why the best people to devise biological experiments are biologists themselves.

1.4 Why good experimental design is particularly important to life scientists

Two key concepts that crop up continually when thinking about experiments (and consequently throughout this book) are **random variation** and **confounding factors**. Indeed it might be said that the two major goals of designing experiments are to minimize random variation and account for confounding factors. Both of these subjects will be covered in depth in Chapters 3 and 4 respectively. However, we mention them here in order to give a flavour of why experimental design is particularly important to life scientists.

1.4.1 Random variation

Random variation

is the differences between measured values of the same variable taken from different experimental subjects.

Random variation is also called **between-individual variation, inter-individual variation, within-treatment variation** or **noise**. This simply quantifies the extent to which individuals in our sample (which could be animals, plants, human subjects, study plots or tissue samples, to give but a few examples) differ from each other. We discuss the consequences of this fully in Chapter 3. For example, all 10-year-old boys are not the same height. If our study aims to study national differences in the height of 10-year-old boys, we should expect that all boys in our study will not have the same height, and that this variation in height will be driven by a large number of factors (e.g. diet, socio-economic group) as well as the possible effect of the factor we are interested in (nationality).

Q 1.3 Humans have tremendous variation in the patterning of grooves on our fingers, allowing us to be individually identified by our fingerprints. Why do you think there is such variation?

Random variation is everywhere in biology. All the sheep in a flock are not exactly the same. Hence, if we want to describe the characteristic weight of a sheep from a particular flock, we generally cannot simply weigh one of the sheep and argue that its weight is also valid for any and all sheep in the flock. It would be better to measure the weights of a representative sample of the flock, allowing us to describe both the average weight of sheep in this flock and the extent of variation around that average. Things can be different in other branches of science: a physicist need only calculate the mass of one electron, because (unlike sheep) all electrons are the same weight. Physicists often need not concern themselves with variation, but for life scientists it is an ever-present concern, and we must take it into account when doing experiments.

Good experiments minimize random variation, so that any variation due to the factors of interest can be detected more easily.

1.4.2 Confounding factors

Definition 1.3

If we want to study the effect of variable *A* on variable *B*, but variable *C* also affects *B*, then *C* is a **confounding factor**.

Often we want to understand the effect of one factor (let's call it variable *A*) on another factor (*B*). However, our ability to do this can be undermined if *B* is also influenced by another factor (*C*). In such circumstances, *C* is called a **confounding factor** (sometimes referred to as a **confounding variable** or **third variable**). For example, if we observe juvenile salmon in a stream to see if there is an effect of the number of hours of sunlight in the day (variable *A*) on rates of foraging (variable *B*), then water temperature (variable *C*) may be a confounding factor. We might expect increased sunlight to increase foraging activity, but we might also expect increased water temperature to have the same effect. Disentangling the influences of temperature and sunlight is likely to be particularly problematic as we might expect sunlight and water temperature to be closely linked, so it will be challenging to understand the effect of one in isolation from the other. Hence confounding factors pose a challenge, but not an insurmountable one, as we'll discuss fully in Chapter 4.

Another advantage that the physicist has is that he or she can often deal with very simple systems. For example, they can isolate the electron that they want to measure in a vacuum chamber containing no other particles that can interact with their particle of interest. The life scientist generally studies complex systems, with many interacting factors. The weight of an individual sheep can change quite considerably over the course of the day through digestion and ingestion, or through being rained on. Hence, something as simple as taking the weight of animals is not as simple as it first appears. Imagine that today you measure the average weight of sheep in one flock and tomorrow you do the same for another flock. Let's further imagine that you find that the average weight is higher for the second flock. Are the animals in the second flock really intrinsically heavier? Because of the way that you designed your experiment it is hard to know. The problem is that you have introduced a confounding factor: time of measurement. The sheep in the two flocks differ in an unintended way; they were measured on different days. If it rained overnight then there might be no real difference between sheep in the two flocks; it might just be that all sheep (in both flocks) are heavier on the second day because they are wetter. We need to find ways to perform experiments that avoid or account for confounding factors, so that we can understand the effects of the factors that we are interested in. Techniques for doing this will be tackled in Chapter 4. However, before we can do this, we have to be clear about the scientific question that we want a given experiment to address. Designing the right experiment to answer a specific question is what the next chapter is all about.

Q 1.4 If we are interested in comparing eyesight between smokers and non-smokers, what other factors could contribute to variation between people in the quality of their eyesight? Are any of the factors that you've chosen likely to influence someone's propensity to smoke?

Q 1.5 Faced with two flocks of sheep 25 km apart, how might you go about measuring sample masses in such a way as to reduce or remove the effect of time as a confounding factor?

Confounding factors make it difficult for us to interpret our results, but their effect can be eliminated or controlled by good design.

Summary

- You cannot be a good life scientist without understanding the basics of experimental design.

- The basics of experimental design amount to a small number of simple rules; you do not have to get involved in complex mathematics in order to design simple but good experiments
- If you design poor experiments, then you will pay in time and resources wasted.
- Time and resource concerns are trivial compared to the imperative of designing good experiments so as to reduce (or hopefully eliminate) costs to your experiment in terms of suffering to animals or humans or disturbance to an ecosystem.