


other people think of them, then they will overestimate other people's positive perceptions. You guessed it, Fester Ingpant-Stain collected yet more data. He asked each contestant to fill out a questionnaire evaluating all of the other contestants' personalities, and also to complete the questionnaire about themselves but answering from the perspective of each of their housemates. (So, for every contestant there is a measure of what they thought of every other contestant, and also a measure of what they believed every other contestant thought of them.) He found out that the contestants with personality disorders did overestimate their housemates' opinions of them; conversely, the contestants without personality disorders had relatively accurate impressions of what others thought of them. These data, irritating as it would be for me, support Fester Ingpant-Stain's theory more than mine: contestants with personality disorders do realize that they have unusual personalities but believe that these characteristics are ones that others would feel positive about. Fester Ingpant-Stain's theory is quite good: it explains the initial observations and brings together a range of research findings. The end result of this whole process (and my career) is that we should be able to make a general statement about the state of the world. In this case we could state 'Reality TV contestants who have personality disorders overestimate how much other people like their personality characteristics'.



Based on what you have read in this section, what qualities do you think a scientific theory should have?

## 1.6 Collecting data: measurement

In looking at the process of generating theories and hypotheses, we have seen the importance of data in testing those hypotheses or deciding between competing theories. This section looks at data collection in more detail. First we'll look at measurement.

### 1.6.1 Independent and dependent variables

To test hypotheses we need to measure variables. Variables are things that can change (or vary); they might vary between people (e.g., IQ, behaviour) or locations (e.g., unemployment) or even time (e.g., mood, profit, number of cancerous cells). Most hypotheses can be expressed in terms of two variables: a

proposed cause and a proposed outcome. For example, if we take the scientific statement, ‘Coca-Cola is an effective spermicide’<sup>8</sup> then the proposed cause is ‘Coca-Cola’ and the proposed effect is dead sperm. Both the cause and the outcome are variables: for the cause we could vary the type of drink, and for the outcome, these drinks will kill different amounts of sperm. The key to testing scientific statements is to measure these two variables.

<sup>8</sup> Actually, there is a long-standing urban myth that a post-coital douche with the contents of a bottle of Coke is an effective contraceptive. Unbelievably, this hypothesis has been tested and Coke does affect sperm motility (movement), and some types of Coke are more effective than others – Diet Coke is best, apparently (Umpierre, Hill & Anderson, 1985). In case you decide to try this method out, I feel it worth mentioning that despite the effects on sperm motility a Coke douche is ineffective at preventing pregnancy.

#### Cramming Sam’s Tips Variables



When doing and reading research you’re likely to encounter these terms:

- *Independent variable*: A variable thought to be the cause of some effect. This term is usually used in experimental research to describe a variable that the experimenter has manipulated.
- *Dependent variable*: A variable thought to be affected by changes in an independent variable. You can think of this variable as an outcome.
- *Predictor variable*: A variable thought to predict an outcome variable. This term is basically another way of saying ‘independent variable’. (Although some people won’t like me saying that; I think life would be easier if we talked only about predictors and outcomes.)
- *Outcome variable*: A variable thought to change as a function of changes in a predictor variable. For the sake of an easy life this term could be synonymous with ‘dependent variable’.



A variable that we think is a cause is known as an **independent variable** (because its value does not depend on any other variables). A variable that we think is an effect is called a **dependent variable** because the value of this variable depends on the cause (independent variable). These terms are very closely tied to experimental methods in which the cause is manipulated by the experimenter (as we will see in [Section 1.7.2](#)). However, researchers can't always manipulate variables (for example, if you wanted see whether smoking causes lung cancer you wouldn't lock a bunch of people in a room for 30 years and force them to smoke). Instead, they sometimes use correlational methods ([Section 1.7](#)), for which it doesn't make sense to talk of dependent and independent variables because all variables are essentially dependent variables. I prefer to use the terms **predictor variable** and **outcome variable** in place of dependent and independent variable. This is not a personal whimsy: in experimental work the cause (independent variable) is a predictor, and the effect (dependent variable) is an outcome, and in correlational work we can talk of one or more (predictor) variables predicting (statistically at least) one or more outcome variables.




## 1.6.2 Levels of measurement


Variables can take on many different forms and levels of sophistication. The relationship between what is being measured and the numbers that represent what is being measured is known as the **level of measurement**. Broadly speaking, variables can be categorical or continuous, and can have different levels of measurement.

A **categorical variable** is made up of categories. A categorical variable that you should be familiar with already is your species (e.g., human, domestic cat, fruit bat, etc.). You are a human or a cat or a fruit bat: you cannot be a bit of a cat and a bit of a bat, and neither a batman nor (despite many fantasies to the contrary) a catwoman exist (not even one in a PVC suit). A categorical variable is one that names distinct entities. In its simplest form it names just two distinct types of things, for example male or female. This is known as a **binary variable**. Other examples of binary variables are being alive or dead, pregnant or not, and responding 'yes' or 'no' to a question. In all cases there are just two categories

and an entity can be placed into only one of the two categories. When two things that are equivalent in some sense are given the same name (or number), but there are more than two possibilities, the variable is said to be a **nominal variable**. It should be obvious that if the variable is made up of names it is pointless to do arithmetic on them (if you multiply a human by a cat, you do not get a hat). However, sometimes numbers are used to denote categories. For example, the numbers worn by players in a sports team. In rugby, the numbers on shirts denote specific field positions, so the number 10 is always worn by the fly-half<sup>9</sup> and the number 2 is always the hooker (the ugly-looking player at the front of the scrum). These numbers do not tell us anything other than what position the player plays. We could equally have shirts with FH and H instead of 10 and 2. A number 10 player is not necessarily better than a number 2 (most managers would not want their fly-half stuck in the front of the scrum!). It is equally daft to try to do arithmetic with nominal scales where the categories are denoted by numbers: the number 10 takes penalty kicks, and if the coach found that his number 10 was injured, he would not get his number 4 to give number 6 a piggy-back and then take the kick. The only way that nominal data can be used is to consider frequencies. For example, we could look at how frequently number 10s score compared to number 4s.

<sup>9</sup> Unlike, for example, NFL football where a quarterback could wear any number from 1 to 19.

Jane Superbrain 1.2 Self-report data 



A lot of self-report data are ordinal. Imagine two judges on *The X Factor* were asked to rate Billie's singing on a 10-point scale. We might be confident that a judge who gives a rating of 10 found Billie more talented than one who gave a rating of 2, but can we be certain that the first judge found her five times more talented than the second? What if both judges gave a rating of 8; could we be sure that

they found her equally talented? Probably not: their ratings will depend on their subjective feelings about what constitutes talent (the quality of singing? showmanship? dancing?). For these reasons, in any situation in which we ask people to rate something subjective (e.g., their preference for a product, their confidence about an answer, how much they have understood some medical instructions) we should probably regard these data as ordinal, although many scientists do not.



So far, the categorical variables we have considered have been unordered (e.g., different brands of Coke with which you're trying to kill sperm), but they can be ordered too (e.g., increasing concentrations of Coke with which you're trying to kill sperm). When categories are ordered, the variable is known as an **ordinal variable**. Ordinal data tell us not only that things have occurred, but also the order in which they occurred. However, these data tell us nothing about the differences between values. In TV shows like *The X Factor*, *American Idol*, and *The Voice*, hopeful singers compete to win a recording contract. They are hugely popular shows, which could (if you take a depressing view) reflect the fact that Western society values 'luck' more than hard work.<sup>10</sup> Imagine that the three winners of a particular *X Factor* series were Billie, Freema and Elizabeth. The names of the winners don't provide any information about where they came in the contest; however, labelling them according to their performance does – first, second and third. These categories are ordered. In using ordered categories we now know that the woman who won was better than the women who came second and third. We still know nothing about the differences between categories, though. We don't, for example, know how much better the winner was than the runners-up: Billie might have been an easy victor, getting many more votes than Freema and Elizabeth, or it might have been a very close contest that she won by only a single vote. Ordinal data, therefore, tell us more than nominal data (they tell us the order in which things happened) but they still do not tell us about the differences between points on a scale.

<sup>10</sup> I am in no way bitter about spending years learning musical instruments and trying to create original music, only to be beaten to musical fame and fortune by 15-year-olds who can sing, sort of.

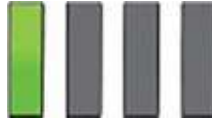
The next level of measurement moves us away from categorical variables and

into continuous variables. A **continuous variable** is one that gives us a score for each person and can take on any value on the measurement scale that we are using. The first type of continuous variable that you might encounter is an **interval variable**. Interval data are considerably more useful than ordinal data, and most of the statistical tests in this book rely on having data measured at this level at least. To say that data are interval, we must be certain that equal intervals on the scale represent equal differences in the property being measured. For example, on [www.ratemyprofessors.com](http://www.ratemyprofessors.com), students are encouraged to rate their lecturers on several dimensions (some of the lecturers' rebuttals of their negative evaluations are worth a look). Each dimension (helpfulness, clarity, etc.) is evaluated using a 5-point scale. For this scale to be interval it must be the case that the difference between helpfulness ratings of 1 and 2 is the same as the difference between (say) 3 and 4, or 4 and 5. Similarly, the difference in helpfulness between ratings of 1 and 3 should be identical to the difference between ratings of 3 and 5. Variables like this that look interval (and are treated as interval) are often ordinal – see [Jane Superbrain Box 1.2](#).

**Ratio variables** go a step further than interval data by requiring that in addition to the measurement scale meeting the requirements of an interval variable, the ratios of values along the scale should be meaningful. For this to be true, the scale must have a true and meaningful zero point. In our lecturer ratings this would mean that a lecturer rated as 4 would be twice as helpful as a lecturer rated with a 2 (who would, in turn, be twice as helpful as a lecturer rated as 1). The time to respond to something is a good example of a ratio variable. When we measure a reaction time, not only is it true that, say, the difference between 300 and 350 ms (a difference of 50 ms) is the same as the difference between 210 and 260 ms or between 422 and 472 ms, but it is also true that distances along the scale are divisible: a reaction time of 200 ms is twice as long as a reaction time of 100 ms and half as long as a reaction time of 400 ms. Time also has a meaningful zero point: 0 ms does mean a complete absence of time.

*Continuous variables* can be, well, continuous (obviously) but also discrete. This is quite a tricky distinction ([Jane Superbrain Box 1.3](#)). A truly continuous variable can be measured to any level of precision, whereas a **discrete variable** can take on only certain values (usually whole numbers) on the scale. What does this actually mean? Well, our example of rating lecturers on a 5-point scale is an example of a discrete variable. The range of the scale is 1–5, but you can enter only values of 1, 2, 3, 4 or 5; you cannot enter a value of 4.32 or 2.18. Although a continuum exists underneath the scale (i.e., a rating of 3.24 makes sense), the actual values that the variable takes on are limited. A continuous variable would be something like age, which can be measured at an infinite level of precision


(you could be 34 years, 7 months, 21 days, 10 hours, 55 minutes, 10 seconds, 100 milliseconds, 63 microseconds, 1 nanosecond old).




### 1.6.3 Measurement error

It's one thing to measure variables, but it's another thing to measure them accurately. Ideally we want our measure to be calibrated such that values have the same meaning over time and across situations. Weight is one example: we would expect to weigh the same amount regardless of who weighs us, or where we take the measurement (assuming it's on Earth and not in an anti-gravity chamber). Sometimes, variables can be measured directly (profit, weight, height) but in other cases we are forced to use indirect measures such as self-report, questionnaires, and computerized tasks (to name a few).

It's been a while since I mentioned sperm, so let's go back to our Coke as a spermicide example. Imagine we took some Coke and some water and added them to two test tubes of sperm. After several minutes, we measured the motility (movement) of the sperm in the two samples and discovered no difference. A few years passed, as you might expect given that Coke and sperm rarely top scientists' research lists, before another scientist, Dr Jack Q. Late, replicated the study. Dr Late found that sperm motility was worse in the Coke sample. There are two measurement-related issues that could explain his success and our failure: (1) Dr Late might have used more Coke in the test tubes (sperm might need a critical mass of Coke before they are affected); (2) Dr Late measured the outcome (motility) differently than us.



Jane Superbrain 1.3 Continuous and discrete variables



The distinction between continuous and discrete variables can be blurred. For one thing, continuous variables can be measured in discrete terms; for example, when we measure age we rarely use nanoseconds but use years (or possibly years and months). In doing