This is copied (with permission) from the book

Why do buses come in threes?

The hidden mathematics of everyday life

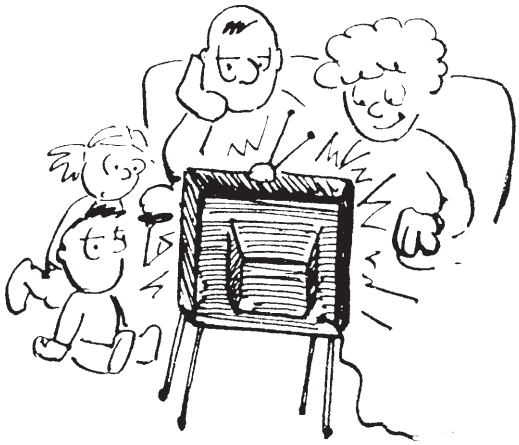
by

Rob Eastaway and Jeremy Wydham

and published by

Robson Books 1998

ISBN 186105 162



3

HOW MANY PEOPLE WATCH CORONATION STREET?

Most public statistics come from surveys,
but how reliable are they?

According to the official figure published in the newspapers, 16.8 million people tuned in to watch *Coronation Street* on 29 September 1997. That is about a quarter of the UK population, a vast number of people.

But hang on a second: how do they know? Did anyone come and ask *you* whether you were watching it? Are there spies peeking in through everyone's windows? Is there some device at ITV for detecting how many TV sets are sucking their signal through the ether? Thankfully, Big Brother is not watching you. ITV knew how many people watched the programme thanks to maths. The number 16.8 million was produced by the mathematics of sampling.

Sampling is the science of estimating how many people are doing something by asking just a few of them. Strictly speaking, sampling cannot guarantee to produce the true answer. If ITV wanted to know *exactly* how many people were watching *Coronation Street* on a particular evening they would have no alternative but to monitor every household in the country. However, the cost of doing so would far exceed their budgets so it is not a serious option. But then, who needs to have the precise answer? If the number watching was really16.6 and not 16.8 million, the programme would still be shown. Many, if not most, of the statistics that we deal with every day don't have to be correct to a high level of presision. Done properly, small samples can produce remarkably accurate estimates.

Taking samples is big business. There are hundreds of research companies in the UK alone finding out what we eat, what we watch, where we travel, and what we think about it all. So how does sampling work?

Instead of asking 56 million people, you can get extremely close to the right answer by asking only a small fraction of that number. At least, the answer will be extremely close as long as the sample is large enough and the sample is made of a suitable cross-section of the population (to avoid bias).
Above all, however, the respondents must be telling the truth . . . .

The mathematics of lying

In most surveys, the respondents have no particular reason to lie. If somebody is asked if they bought a tin of baked beans in the last week, their response will probably be honest, even if it might be inaccurate - the memory can, after all, play strange games.

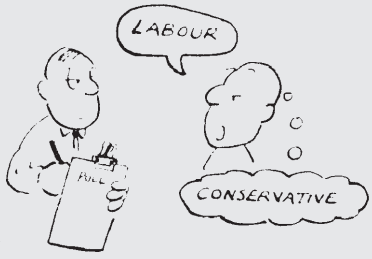
However, in other situations, lying is an issue which pollsters ignore at their peril. Always be suspicious of the answer anybody gives if they are asked for information about their income, sexual activity or, as it has emerged in recent years, politics. Questions on voting were famously the cause of embarrasment to a number of market research agencies in the 1992 General Election. At the start of the election broadcast, the results of opinion polls during the lead up to Election day were used to forecast with some confidence that Labour would win the election, but with a hung parliament.

Little did they know, however, that the survey had overlooked a crucial distortion to the figures. It appears that while Labour voters had been happy to reveal who they had supported, many Conservatives were less willing to admit the truth. Tories felt that admitting to voting Tory might make them sound selfish or greedy. Some Tory voters said they would vote

Election lies? The evidence.

Every single opinion poll conducted prior to the 1992 General Election underestimated the Conservative vote by at least 4.5 per cent. Even the NOP/BBC exit poll, which was based on people who had already cast their vote, produced an underestimate. The results were:

	Con	*Lab*	*Lib-Dem*
NOP/BBC	*40%*	*36%*	*18%*
Actual Result	*43%*	*35%*	*18%*



The sampling technique used meant that the Conservative vote should have been between 37 per cent and 43 per cent. The actual result barely scraped into the range. This leads to one of three conclusions. The pollsters were unlucky in the sample they chose; the sample was not a fair representation of the public; or the individuals were not telling the truth.

Labour, and others simply refused to answer the pollsters. All of this meant that the reported figures were wrong, so wrong in fact that when the final result of the election was known, not only had the Tories won the election but they had won by a clear majority.

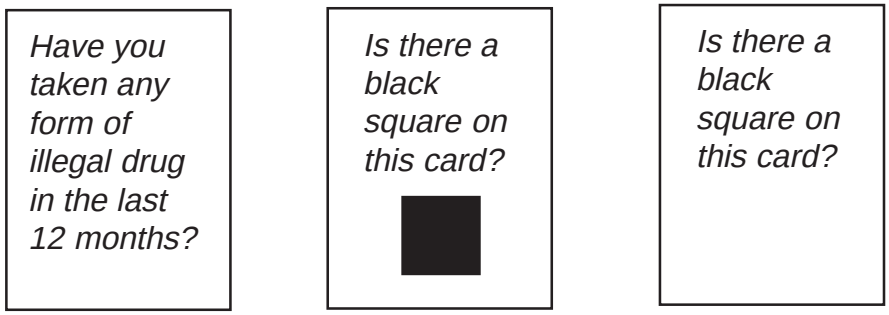
Not all lies are intended to deceive others however. Sometimes people lie in order to deceive themselves because admitting the truth to oneself can be painful. This can apply in television viewing, for example. A member of the public who discovers he watched 35 hours of TV last week may not be prepared to admit to himself that he is *such* a couch potato, and so he ticks the box labelled 21-30 hours.

One amusing example of where researchers could identify mathematically that the answers they were receiving were not truthful was a survey many years ago on the sexual habits of men and women. As part of this survey, the respondents were asked how many different members of the opposite sex they had ever slept with. The average answer for males was 3.7. The average answer for females was 1.9. Now since this survey was taken over a large and representative sample, the answers to the questions should have been the same. After all, if a man sleeps with a woman, then the woman is also sleeping with the man, and so one is added to each tally. The conclusion that the researcher came to was that men tend exaggerate the number of sleeping partners that they have had, while women prefer to trim the figures a little . . . [1]

Statisticians need to devise suitable techniques to identify, and then eliminate, the distortion caused by lying.

Some mathematical tricks have been devised in order to help researchers obtain answers to awkward questions. During the Vietnam war, the American authorities needed to know how many of the troops were taking drugs. Drug-taking was rumoured to be high, and it was important to establish whether this was true. However, no soldier with any sense wanted to admit to taking drugs, a criminal offence. So how could the researchers get to the truth? They used a technique similar to what follows.

The researcher has a bag which contains three pieces of card, which he shows to the soldier. The three cards are:



1 This not the only possible interpretation - for example, it could be that *both* parties exaggerate or underestimate. Another interpretation is that women find the experience more forgettable.

The soldier put his hand into the bag and removes a card at random without showing it to the researcher. He then ticks Yes or No on the researcher's answer sheet.

If he ticks Yes, it is either because he picked the card with the black square on it, or because he picked the drugs question and he is admitting that he has taken drugs. The researcher cannot know which it is, and so the soldier cannot be incriminated, which means he should be more inclined to be honest.

Now here is the clever part. Suppose that the researcher interviews 1,200 soldiers in this way, and that at the end of the survey 560 have answered yes to the question on their card. On average, 400 of them picked out the card with the square on it, 400 selected the card without the square and 400 picked the drugs card. This means that ot the 560 yes's, about 400 were responses to the square question, which leaves 160 which drug answers. The best estimate, therefore, is that 160 out of 400 soldiers took drugs, or 40 per cent.

This is a simplification of what a market researcher would do, and we have made up the numbers here. Nevertheless, this type of survey was conducted on the US soldiers and it emerged that many US soldiers did indeed take illegal drugs during the war.

Have enough people been asked?

'*In a recent test, 80 per cent of cat owners said that their cats preferred Furry Paws biscuits.*' It sounds impressive, and if you are a cat owner, you might be inclined to give the product a try if you see it on the shelf. But if you now discover that this 'test' actually involved only ten cat owners, you ought to be a lot less impressed.

It is stretching belief to think that in every group of ten owners there will be exactly eight fans of Furry Paws. In fact if the researchers repeated this test, the results would keep changiing. They might get the results 20 per cent, 50 per cent, 30 per cent, 0 per cent, 80 per cent. The last result would allow them to truthfully say that 'in one recent test 80 per cent preferred Furry Paws'.

It should come as no surprise that the more people that are surveyed, the more likely it is that the answer is close to the correct one. A survey of 100 people should be more accurate than one of 10, and 1,000 will be more accurate still. Taken to its limit, if the entire population is interviewed that the result is certain to be exactly right.

How large a sample do you need to take before you have enough? That depends on what you mean by enough, and it also depends on what you are testing for. In most everyday surveyus, like opinion polls or tests for how many people have seen an advertisement, asking 1,000 pelple is usually sufficient to give a result accurate to within 5%.

However, there are exceptions. In the 1930s, the US authorities wanted to check the effectiveness of a polio vaccine, so 450 children were innoculated with the vaccine. 680 who had not been vaccinated (and who came from the same background as the test group) were monitored as a control group. Soon afterwards there was a serious outbreak of polio. None of the 450 who had been vaccinted caught the disease. Nor did any of the 68o who

were unprotected. As a result, the experiment proved absolutely nothing. Even during a serious outbreak, the infection rate of polio is so low that researchers would have needed a sample of thousands before it became likely that the control group would include polio cases, so that the experiment would have meaningful results.

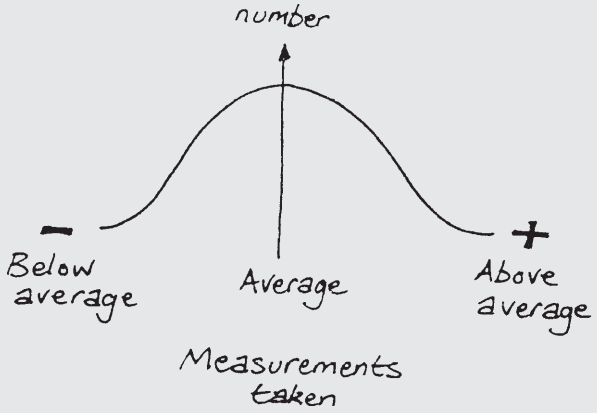
Statisticians have a precise way of stating how confident they are in a result of a survey. Suppose in the cat food survey, 80 per cent say their cat prefers Furry Paws. The correct statistical way to present this result is as a range rather than as a precise figure. If 1,000 cat owners had been surveyed, then the statistician would say:

"The true figure is in the range 77 per cent to 83 per cent with a 95 per cent confidence"

This statement is easy to misinterpret. What the words in the quote mean is: "We are likely to be right by stating that the true value is somewhere between 77 per cent and 83 per cent - but there is a one in twenty chance that the answer is outside that range."

Telescopes and sampling errors

The mathematician Gauss (1777-1855) was also a keen astronomer. He acquired a new telescope, and decided to use it to produce a more accurate calculation of the diameter of the moon. To his surprise, he discovered that every time he took a measurement, his answer was slightly different. He plotted the results and found that they formed a bell shaped curve, with most results close to the central average but the occasional one quite inaccurate.



Gauss quickly realised that any measurement he took was a 'sample' prone to error but which could be used as an estimate of the correct answer. The more readings he took, the closer the average would be to the correct reading. He established that errors in readings belonged to a curve whose complex formula also includes π *and* e*. There they are again!*

The public has grown a lot wiser to some of the tricks os using small samples to ‘prove’ a result, but it still goes on. Management consultancy firms regularly conduct survveys on trends in current business, and put out press releases making claims such as ‘70 per cent of compainies think exports are the key to success’.

Given the small samples they take in these surveys, it would be more honest if they said ‘We are confident that between 50 per cent and 90 per cent of companies think exports are the key to success’. But of course no newspaper would carry a story as boring as that.

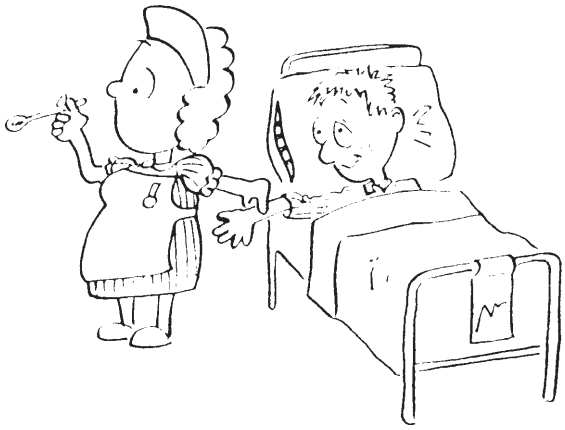
Has the average person been sampled?

The survey has been conducted. A large sample has been taken, and clever techniques have been used to ensure that the answers received have been truthful. Unfortunately this is still not enough to be certain that the results of the survey are accurate. The sample taken has to be *representative*of the whole population too.

One of the large market research companies was commissioned to test out the public's reaction to a new product: tins of baked beans with pork sausages. They selected a district of London which was conveniently located, had a representative income level and a mixed age-group. It was unbiased in every way except one. The district happened to be Golders Green. Golders Green has a very high Jewish population and a correspondingly low appreciation of pork sausages.

It doesn't matter how large the sample is, if it is biased then that bias will never disappear no matter how many people are interviewed. One of the great skills of market research is finding a sample that represents the population as a whole.

One popular way of picking a random sample is by using a telephone directory and picking out every 100th person in the list. This is a cheap way of sampling and it would be perfectly all right, for example, to use a telephone survey to find out which brand of cereal is most popular among the general public. However, it would be be much less advisable to use such a survey to find out what jobs people have. Who is more likely to answer the telephone at a private address, a full time mother or a city lawyer? City lawyers’ hours are probably so long that your chance of finding them at home is almost zero. And there are some professions which are seriosly under-represented in the phone book. To take an extreme example, how many television presenters put their names in the directory?



Bias can crop up in all sorts of unexpected areas. A nurse often takes a patient's pulse rate for twenty seconds and scales this up to give a rate over a minute. The nurse is in fact taking a sample, and it may well be that the sample was not representative of the patient's normal state. An attractive female nurse clutching the wrist of a healthy young male can produce an extremely distorted pulse rate, especially for the first twenty seconds. If the patient is of a nervous disposition and has just been told there may be something wrong with him, this may also lead to a biased result.

Another survey that is vulnerable to bias is the one used to create the pop music charts. Perhaps you weren's even aware that the record charts are the product of a sample. Instead of monitoring the sales of a CD in every store in the country, the chart compilers nominate a number of stores to become part of their survey. The number of discs sold in these stores is added up, and the national figure is produced by scaling up the results from the sample.

Not surprisingly, the chart stores are sworn to secrecy, because if the recording companies knew which stores were used for the charts, they would immediately send out their staff to purchase discs from those stores. This would escalate their record up the charts, thus generating more coverage, and from more coverage come genuine sales. Hype is almost everything in the pop world.

That is why (it is alleged) there is a big espionage game played by the record companies to try to find out which shops are used to compile the charts. One recording company used a particularly clever if devious trick which was to pretend to be conducting market research on record shops.

They asked, ‘Are you happy with the way that the chart compilers currently obtain their information from the shops?’. If the shop had no opinion or didn't understand the question, it meant that they were not involved in the charts. If they did express an opinion, good or bad, or said ‘we aren't allowed to comment’ it demonstrated that they knew something about what was involved and were almost certainly a chart store. In answering one question the stores were inadvertently answering another more important one!

Full list of the chapters in the Book