Session 2

Chapter 2, Sections 2.2 –2.8

2.1 Imagine that a study did find that people who preferred butter were better drivers than those that preferred margarine. Can you think of any hypotheses that could explain this finding, and consider which of these you consider most plausible.

It seems very unlikely that this dietary preference influences driving ability directly. Butter or margarine makes up a very small part of our diet and contains little (in terms of nutrients or trace elements) that is not readily obtained from other commonly consumed foodstuffs. It is possible, but unlikely, that something present in butter but not in margarine has a beneficial effect on eyesight, alertness or reactions. Similarly it is possible but unlikely that something present in margarine but not in butter has a detrimental effect on some aspect of physiological function that degrades driving ability.

Slightly more likely, but still unlikely, is that preference for butter or margarine is an indication of a wider difference in dietary preference between people, and this wider preference affects driving ability directly. For example, it could be that those that prefer margarine have a general preference for refined heavily processed food products over unprocessed foodstuffs. This could lead to a dietary deficiency that does lead to a decrease in say alertness and so driving ability. This is more plausible than our first explanation, but still relatively unlikely compared to the answer below.

The most plausible explanation is a third-variable effect. That is, that preference for margarine or butter has no effect whatsoever on driving ability, but is linked to some other variable that affects driving ability. For example, women might be more likely that men to prefer butter to margarine and might be better drivers than men. Here the third variable is sex, which is related to both the dietary preference and driving ability. Other plausible third variables are age and socio-economic group.

2.2 How would you the measure driving ability in a study like the one described above?

If you are in a country where driving offences are recorded as penalty points on the driving licence, then the quickest and cheapest way would be to record the number of points on a person's driving license. This is inaccurate since a person's probability of having penalty points will be correlated with some but not all aspects of poor driving. Worse, it will be affected by the length of time for which they have held a licence, the amount of driving they do, and the type of diving they do (e.g. business or pleasure). Lastly, penalty points can in substantial part be a matter of good or bad luck.

Better would be to make each participant sit a driving test. This would be considerably more work. It would also be better to ask the examiner to provide a quantitative score of ability rather than just a pass or fail. One person could fail because they cope poorly with a traffic situation that another person with identical driving ability does not experience on their test. This heterogeneity in driving situations experienced during tests by the different people leading to heterogeneity of scores is a problem. One way to cope with this would be to make every candidate sit a number of tests, but this is lots of hard work. Better

might be to use a simulator (now widely available through, for example, BSM) so that each candidate experiences identical sets of (simulated) traffic conditions. Again we could get a qualified driving test examiner to assess performance, rather than attempting to do this ourselves.

2.3 Ecologists commonly use indirect measures such as clutch size, feeding rate, and mass per unit length as indirect measures of fitness. What are the limitations of this, and why do ecologists persist in using these indirect measures in the face of these limitations?

Fitness is very challenging to measure directly. To compare the fitness of two female gannets, we would want to record the number of female offspring that each female produces during her lifetime that survive long enough to reproduce themselves. Given that these birds can live for a decade and that offspring will disperse over hundreds if not thousands of kilometres, measuring fitness directly would be very, very challenging. Hence the attraction of indirect measures.

Now, all other things being equal, if we find that in a given year the mean clutch produced by females of phenotype A is greater than the mean clutch produced by individuals of phenotype B, then this should lead to greater fitness for the genes carried by phenotype A. How good the indirect measure is depends on how reasonable the assumption that all other things being equal is. In our illustrative example, if there is a trade-off between annual clutch produced by an individual and its probability of surviving the winter, then our conclusion than maximising clutch size maximises fitness could be wrong. Hence, we use indirect measures of fitness because fitness is very hard to measure, however we must use our understanding of the general biology of a system to consider how reliable indirect measures of fitness are likely to be.

2.4 Can you think of a scientific study where ethical considerations might drive you to using indirect measurements?

If we wanted to study how increasing amounts of alcohol in someone's bloodstream affects driving ability then it would be ethical to measure driving performance using a simulator rather than a real car.

2.5 In the tail length experiment, we want to have a control group in which tail length is unaltered. Why then do we bother to cut the tails off then glue them back on in exactly the same position?

Other than tail length, we want as few things to be different between the three groups of birds as possible. It is just possible that the trauma of having the tail removed, or our ability to reattach the tail properly has an effect on the subsequent behaviour of the male birds. It is even just possible, even if unlikely, that the smell of the glue influences female mate choice. By taking the trouble to perform these manipulations on all groups, we control for any of these possible confounding factors.

2.6 Discuss how you would test the hypothesis "*Female humans find blue eyes more attractive than brown*" by correlational and manipulative means. Discuss the pros and cons of each, and which you would adopt to address this question.

Correlation: We set up a cocktail party with 20 men, of similar age and occupation, ten with brown eyes and ten with blue. We also invite 20 girls, each independently tasked with identifying the five most attractive men in the room. We then explore whether more blue-eyed than brown-eyed men were selected. Alternatively, this process could be done with a pack of passport-style photographs, asking girls to score the attractiveness of each. This is quicker and easier, and removes added noise due to personality differences, height, build, clothing etc.

Manipulation: In the cocktail party set-up, we could run several parties, with each man wearing clear contact lens to some parties and coloured ones (designed to make them appear as if they have the opposite eye colour) to others. Different girls are invited to each party, but we can explore whether the manipulation makes men appear more or less attractive. With the photographs, we would use image analysis software to change the eye colour. A given individual's photograph would be in each girl's pack of photos, but sometimes he will have brown eyes and sometimes blue.

The cocktail party set-up is a great deal of hard work and introduces possible confounders, so we would do it using photos. The concern we have is potential third variables. For example, that women actually find dark hair attractive and dark hair is correlated with blue eyes. This concern would drive us towards the manipulative approach, but we would need to be very careful that we doctored photos carefully so as to produce realistic looking eyes. It would probably be worthwhile running a pilot project to see if women can detect whether or not a photo has been doctored. If we cannot produce a biologically realistic manipulate then we will have to go for the correlational approach. We could record potential third variables such as hair colour and look for correlations between these and eye colour that might give rise to third variable effects.

2.7 The book suggests that women who go to university are less likely to marry than those that do not. However, the book argues that we should not conclude from this that studying at university in itself causes a reduction in a woman's propensity to get married. Explain this reasoning in your own words.

The initial observation comes from an unmanipulated correlational study, and it may be that a third factor can explain the apparent link between going to university and propensity to marry. For example, it may be that individuals from higher socio-economic groups are both more likely to go to university and (quite separately) less likely to marry. Other potential third factors are ethnic group, religion, or geographic region.

2.8 A driver in their twenties is three times more likely to be involved in a road traffic accident than a driver in their sixties. One explanation for this could be that people become safer drivers as they get older. Can you think of any likely third-variable effects that could provide an alternative explanation of this observation?

The older age group will include large numbers of retired people whose driving is done mainly for pleasure; this will be much less true for the younger age group, a significant fraction of whose driving will be done for work purposes. This may mean that younger drivers spend a larger fraction of their time driving in rush-hour traffic or in poor weather

conditions (when your boss would expect you at work, but you would not choose to go out for a pleasure trip). Further, business driving will often involve driving to somewhere that has to be reached by an appointed time, a pressure that is less common for pleasure trips. Further, it is possible that younger drivers simply drive more miles in a year than older drivers.

2.9 Can you think of an alternative explanation for the observation of the last question, in terms of reverse causation?

Simply, it may not be that driving ability improves with age, but rather that only good drivers live to old age. Under this argument, a given individual's driving ability could remain unchanged through life but we would still see a change in the mean ability across the population with age because poorer drivers are more likely to die in road traffic accidents or get banned from driving and so contribute to the population of younger drivers more than they contribute to the population of older drivers.

2.10. Which of the explanations above do you think is the most important factor explaining the three-fold difference in accident rates between these age groups?

The reverse-causation effect seems unlikely to be able to explain such a large difference. For example, in the UK 4000 people a year, at most, die in road traffic accidents, and only the minority of these will be drivers. Perhaps the number of people who kill themselves, get banned, or scare themselves so badly that they voluntarily give up driving through their bad driving is in the order of 5000 a year. This is a very small fraction of the perhaps 40 million drivers in Britain. Hence, this seems unlikely to provide an explanation. It is more likely that some combination of the conventional explanation and the third variables explain the effects. My guess is that the effect of differences in type of driving (business or pleasure) is likely to be the most important third variable. Although it is quite possible that people become safer drivers as they get older, it would seem surprising if they become safer to a sufficient extent to explain a three fold reduction in annual risk without third variables being involved to some extent.