1 One strength of using event sampling rather than time sampling is that you can make sure you record all the events of particular interest. For example, you are observing parents taking their children to primary school. You want to find out whether they park or stop on the zig-zag lines and what proportion do so. If you use event sampling, you make a record every time someone stops. However, if you use time sampling (e.g. every minute), then you would quite possibly miss many occasions of people parking illegally.

2 One strength of using time sampling instead of event sampling is that it can prevent an observation from becoming biased. For example, instead of looking for a particular behaviour (as with event sampling), you record whatever is happening at the time in question (e.g. every 60 seconds). This can mean that you record a frequent behaviour that you had not been aware of as being a regular occurrence (or hadn’t even noticed it) when doing time sampling.

3 a Could do either. As the focus seems to be very much on the different activities, then event sampling might be best to find out the frequency of each time someone plays with the sandpit versus plays with the drawing/painting equipment etc. The observers (probably best to have more than one), could each monitor a particular set of activities and make a tally each time a different child engages with the activity.

b ‘Observing what dog owners do …’ – this sounds a bit more open-ended and it may be the researchers have little idea what to put on an event sampling list. Therefore, time sampling might be more appropriate. For example, an observer might watch one dog walker at a time (possibly using CCTV, or by being stationed at a location in the park with a good viewpoint). For each dog walker, the observer could note down every 30 seconds what he/she is doing – e.g. walking with the dog on the lead, throwing sticks for the dog, talking to other people, picking up dog poo, shouting for the dog etc. A number of dog walkers could be observed in the same way.

c Event sampling would be most appropriate here as the focus is on the products which are being bought. An observer could stand observing a checkout. Every time a product is scanned, he or she could make a note of what it is. There could be several observers watching several checkouts in the supermarket.

4 One strength of structured observations is that they are more objective. Because the behaviour categories and checklists are set up in advance, and the observations pre-structured of what is going to be observed, all of those behaviours will be recorded (whereas, they might be overlooked in an unstructured observation which can be more impressionistic, and just follow the attention or interest of the observer. Another advantage of a structured observation is that it is better than an unstructured observation for producing quantitative data – so that it is possible to compare the frequencies of observations in the different categories of behaviour, and this enables you to find out which are most and least frequent.

5 a (i) shouting encouraging things at their team, (ii) shouting negative things at their team, (iii) singing.

b I would use time sampling and observe e.g. 10 people every 1 minute and record their behaviour at that time and at the same interval for the duration of the match.

c They could be observed through CCTV. Cameras are often located discreetly and people are fairly used to the presence of CCTV cameras so they should not affect people’s behaviours.

d (i) People do not know they are being observed and so informed consent might be an issue. (ii) Another issue is privacy – being observed at close quarters may feel like an invasion of privacy. (iii) Debriefing might be an issue too – participants will not know what data about them has been collected or for what purposes it might be used.

e For the issues of informed consent and privacy, nothing needs to be done. This is because people are in a public place and know that their behaviour is not private and ordinarily observable by strangers. Therefore, this is not a problem according to the BPS code of ethics. For a debrief, participants could be approached at the end of the match and told about the study and its purpose. They could be reassured about anonymity and given the opportunity to withdraw their data at that point or given some contact details so they can contact the researchers at any point in the future if they change their mind.

6 a (i) whether the people are eating, (ii) engaging in prosocial behaviour (e.g. chatting, standing in a queue, putting rubbish/trays away), (iii) engaging in antisocial behaviour (e.g. pushing in to the front of a queue, leaving rubbish behind etc.)

b A random sample of 10 people could be observed for the duration of their time in the canteen. Every behaviour will be marked in a behaviour checklist (see part a) – this is event sampling.

c Either use CCTV again. Or, another student could be the observer and keep a discreet note of all the behaviours of the target person.

d Same for 5 d.

e Same for 5 e.

Qs (page 7)

Study A

1 One suggested procedure: the researcher would, on several days, observe the tigers for one hour before and after zoo opening hours and two hours during zoo opening hours. They would not observe the tigers during feeding times as this might confound or confuse the results. The researcher would observe the tigers using an event sampling procedure with a behaviour checklist. The behaviours on the checklist could include: (i) sleeping, (ii) relaxing, (iii) pacing up and down, (iv) displaying aggressive behaviour (but not towards another tiger) such as growling, (v) interacting positively with other tigers, (vi) interacting negatively with other tigers. The different categories of behaviour could be compared for the two scenarios (zoo open to public versus zoo closed to public).

2 One strength of an observational technique for this study is that it is the most obvious, natural and valid technique for this issue – to find out whether the tigers are affected by the presence of people. Because the focus is on the tigers’ behaviour, what better way is there to find out than by actually watching them?

One weakness of the observational technique is that it is very intensive and time consuming. Observers have to be very astute so that they note down every behaviour of the tigers and do not miss anything. They also have to make decisions quite rapidly about any observed behaviour and about what behaviour category it might fit into – and that might compromise the validity.

3 A different way to conduct the study would be to use recording equipment. This means that the observer(s) can re-watch behaviours to double check the behaviour categorisation.

4 Inter-rater reliability refers to the consistency of codings between two or more observers. If inter-rater reliability is high, then there is high agreement between the observers.

5 Two observers could independently watch the tigers at the same time (or the same recording) and code the same behaviours. Then, later, they would see how much agreement there was between the two observers for each behaviour observed and categorised. If there is high agreement (i.e. more than 80% of total observations in agreement) then that means there is high inter-rater reliability.

6 One threat to validity might be that the time of day affects the tigers’ behaviour, rather than the presence of the public.
One way to deal with this would be to observe the tigers at the same time each day but on days when the zoo is closed, and compare the observed behaviour at the same time of day when the zoo is open, rather than watching them before, during and after the opening hours.

Study B

1 Observers could video doctors' consultations with patients. It would be best to use video (with permission) rather than direct observation because otherwise patients and doctors might alter their behaviour if a third person is sitting in (possibly quite a small) room. Observers could categorise and measure a number of things including the amount of time spent (i) overall for the consultation, (ii) with the doctor talking/explaining, (iii) the doctor asking questions, (iv) the patient talking/explaining, (v) the patient asking questions, (vi) the doctor physically examining the patient, e.g. taking temperature, blood pressure etc., (vii) eye contact with patient versus writing notes, looking at computer etc. They could also rate the doctor for concern/empathy shown.

2 One strength of observation here is that it is likely to be more valid than other measures e.g. questionnaire. This is because in a questionnaire, doctors may alter their responses to make doctors appear in a good light, or because they don’t realise exactly how they behave. In an observation, you can actually see how the doctors behave. However, with an observation where the doctors know they are being observed, they might change their behaviour so they appear the best way. (However, it is harder to maintain good appearances when being observed, than when just simply answering a questionnaire).

3 A different way to conduct this would be to deceive doctors so that they did not know they were being observed/recorded – although this would be highly unethical. You could ask their permission retrospectively though doctors may still feel it was unethical for anyone to observe a confidential consultation without permission.

4 See above.

5 Two researchers could watch a selection of doctor–patient consultations. For each one, they could see whether they agree upon the observed time for each behaviour category (such as the doctor asking questions). For each instance, the observations in each category could be compared. If there is good inter-rater agreement, this would be indicated by a high positive correlation.

6 One threat to the validity of this observation might be that it would be very difficult to get accurate measurements for the length of time for each category. That is because conversations are actually quite messy – people interrupt, two people talk at the same time and so on. This makes it difficult to time observations accurately.

7 One way to deal with this particular problem would be for the observer to have an additional category to time, which would be ‘both patient and doctor talking at the same time/rapid turn-taking’.

Qs (page 9)

1 One strength of this question is that it is open-ended and so respondents can choose to respond in any way they want and not be constrained by a predetermined list. This should increase the validity.

One weakness of this question is that respondents might still feel they have to answer in a certain way because of demand characteristics or because of social desirability. They might feel they have to put down the name of one of the researchers if he or she is a close friend; or that to put down ‘my grandma’ might be a bit uncool, even though it is true. Another weakness is that the question might be overly simplistic. ‘Love’ can take many forms – the love for your baby brother and the love for your boyfriend are very different and it might be difficult for participants to answer that kind of question validly (meaning fully).

2 The question is open-ended because the questionnaire does not ask them to choose from particular pre-set responses. The respondents can write down anything they want from ‘step-grandmother’ to ‘the boy-next-door’.

3 ‘I have a close and trusting relationship with my mother – Strongly agree/agree/not sure/disagree/strongly disagree.’

4 One strength of this question is that the wording is very clear and unambiguous.

One weakness of this question is that, if there is some problem between the adolescent and the mother, such as they have a very difficult relationship, or even that the mother is absent, then the question could be quite upsetting.

5 One way this could be done is to have a list of likely relationships and ask them about how close and trusting they think the relationship is, such as: Rate the following people for how close and trusting your relationship is:

- Your mother
- Your father
- Sibling #1
- Sibling #2
- Sibling #3
- Your best friend
- Your girl/boyfriend

6 One ethical issue that might arise is that of invading a person’s privacy. Asking questions about people’s relationships is quite personal and while many participants may be perfectly happy to be quite open about it, some may find it quite intrusive, especially where a relationship is difficult, complex or even secret/taboo. One way that the student researchers could deal with this would be to tell respondents that they should not feel that they have to answer all questions, especially if they find a question too sensitive. Also, they should be assured of their right to withdraw at any time (even after data has been collected) and told that all results will be treated confidentially and presented anonymously.

7 One other way of collecting information about relationships, other than a questionnaire, would be an interview method. Interviewers could use a structured interview and include a mixture of open and closed questions. This would be face to face and the questions would be predetermined.

8 Respondents might give more detailed responses to the open questions because it takes less effort perhaps to talk than to write things down. However, people might, in some cases, be less ‘open’ about some things because if something is quite personal it could be more difficult to disclose it to an actual human being (as in an interview) rather than ‘anonymously’ a questionnaire. Therefore, this might change the results and in some cases lead to more social desirability, e.g. report increased levels of closeness or love for parents because some people may not want to admit it if they don’t like or don’t get on with parents.

Qs (page 11)

1 A good sampling method for this study would be an opportunity sample, i.e. asking people who are available to you to answer the questionnaire.

2 An opportunity sample in this study could be obtained by asking people, such as friends in the same class or in the college canteen, to take part in the study.

3 One strength of the opportunity sample is that it is easy to find and locate people – it is probably the least problematic way of gaining a sample. One weakness is that the sample is unlikely to be representative of all teenagers – or even of all teenagers in your college. This is because you are likely to pick friends and people who are ‘like you’ and your social groups rather than not like you/not in your social groups.
SUGGESTED ANSWERS

Chapter 1 Psychological investigations

4 One way to overcome this would be to use a random sample of teenagers in your college. This could be achieved by putting everyone’s names on the school roll into a hat and picking e.g. the first 20 names. Every person in the college then has an equal chance of participating. In theory, a random sample is more likely to achieve a representative sample than an opportunity sample.

5 Table: Summary of results

<table>
<thead>
<tr>
<th>Person respondent is closest to</th>
<th>Percentage of respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boyfriend/girlfriend</td>
<td>34%</td>
</tr>
<tr>
<td>Best friend</td>
<td>28%</td>
</tr>
<tr>
<td>Mother</td>
<td>15%</td>
</tr>
<tr>
<td>Father</td>
<td>2%</td>
</tr>
<tr>
<td>Brother/sister</td>
<td>11%</td>
</tr>
</tbody>
</table>

b Percentage of respondents (%)

<table>
<thead>
<tr>
<th>Percentage of respondents (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
</tr>
<tr>
<td>35</td>
</tr>
<tr>
<td>30</td>
</tr>
<tr>
<td>25</td>
</tr>
<tr>
<td>20</td>
</tr>
<tr>
<td>15</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>0</td>
</tr>
</tbody>
</table>

3 a An interview would involve a person – an adult – asking questions. A child may find this intimidating perhaps and children as participants would be more prone to interviewer’s expectations. This would affect the validity of their answers. Therefore, in this respect, a questionnaire might be preferable.

b A questionnaire might be problematic for children aged 8–10 years because their reading and writing skills may not be up to scratch. They might not be able to read all the words, they may not want to write full answers because writing is a relatively tiring business, they may not be able to express themselves adequately in writing, and there may be some problems in understanding the instructions on the questionnaire. So, for this reason, an interview might be preferable because someone is asking the questions (effectively, reading them out loud) and can, if necessary, explain or rephrase them to increase understanding, and it removes the burden of physically writing from the child.

c Test–retest method. The same questionnaire could be given to the same children some time later, e.g. 2 weeks later. The two sets of scores can be compared using a correlation coefficient. If there is high reliability, this means that, overall, each child has got a similar score on both occasions (i.e. it is consistent over time) and this will give a high positive correlation between their scores on both occasions.

d The researcher would need to look at the wording of the questions carefully because it might be that some of the questions are ambiguous or difficult to interpret and this is why the children are giving different answers on different occasions. The researcher could plot the questionnaire with children to find out if there are particularly confusing questions and/or get other researchers to check the questionnaire.

e One way to do this is give children another questionnaire which also claims to be measuring the same thing, i.e. internet usage. If the questionnaire is valid, each child should, overall, get similar scores on both questionnaires. This is called concurrent validity.

4 a Eating disorders are quite a sensitive area. Many people who suffer from eating disorders actually try to disguise their behaviours. Therefore, it may be unlikely that researchers would get a true picture of people’s experiences if they ask them face-to-face. People sometimes feel that they can reveal more personal data if they can write it down (as in a questionnaire) and do not have to worry about someone’s reaction to it.

b A stratified sample would be the best. The researcher would want an equal number of males and females (e.g. 30 of each) and these comprise the subgroups or strata. In the researcher’s city, they could use the electoral register of all 17-year-olds and from this, randomly select 30 males and 30 females. This would be the best way as it would produce the sample most representative of the target population and not be biased towards friends or a particular college.

Qs (page 12)

1 This means that the questionnaire does not give consistent results, so that if the questionnaire was repeated with the same person some time later (e.g. a couple of weeks), then their answers might give different results.

2 One advantage of using self-report measures is that it allows researchers to find out things about people’s personal attitudes or beliefs or private lives – i.e. things which are difficult or impossible to observe. Another advantage of using self-report measures is that they are generally very quick and easy to use and collect data from a large number of people therefore producing lots of data with the least amount of effort (when compared with e.g. experiments or observations).

One disadvantage of self-report measures is that people often give answers which are socially desirable rather than really true. For example, they might give an answer in an interview or questionnaire which makes them appear more interesting, more honest, more thoughtful etc. This will affect the validity of the self-report.

Another disadvantage of self-report measures is that people are not necessarily aware of what they really do and how they behave or, if asked about events in the past, that they cannot remember them accurately. This means that even when people are trying to answer as honestly as possible, the data may not be as valid as, say, an observation which involves recording when people actually do.

3 a An interview would involve a person – an adult – asking questions. A child may find this intimidating perhaps and children as participants would be more prone to interviewer’s expectations. This would affect the validity of their answers. Therefore, in this respect, a questionnaire might be preferable.

b A questionnaire might be problematic for children aged 8–10 years because their reading and writing skills may not be up to scratch. They might not be able to read all the words, they may not want to write full answers because writing is a relatively tiring business, they may not be able to express themselves adequately in writing, and there may be some problems in understanding the instructions on the questionnaire. So, for this reason, an interview might be preferable because someone is asking the questions (effectively, reading them out loud) and can, if necessary, explain or rephrase them to increase understanding, and it removes the burden of physically writing from the child.

c One conclusion is that adolescents are most likely to say that they feel closest to their boyfriend/girlfriend. Another conclusion is that they are least likely to say that feel closest to their father.

Qs (page 14)

Study: weather and mood

1 IV is the weather. The two conditions could be operationalised as follows. Good weather = a day when the temperature is > 22 degrees Celsius, the sun is shining with very little cloud cover. Bad weather = a day when the temperature is < 17 degrees Celsius, there is mainly cloud cover and it is raining. The DV in this study is mood. It is quite difficult to observe a person’s mood (you would have to judge their mood from their behaviours). Therefore, it may be unlikely that researchers would get a true picture of people’s experiences if they ask them face-to-face. People sometimes feel that they can reveal more personal data if they can write it down (as in a questionnaire) and do not have to worry about someone’s reaction to it.

2 People get a higher score on the mood checklist (where a high score indicates positive mood) on days of good weather (sunny with temperature > 22 degrees Celsius) when compared with days of bad weather (cloudy with temperature < 17 degrees Celsius).
Qs (page 17)

1 a IV is whether the child is an only child or has siblings. This is easy to operationalise and people could be categorised on the basis of their answer to a question 'Do you have siblings – yes or no?' The DV is how well children do at school. This could be operationalised as mean GCSE score – where the score is calculated by converting each GCSE grade to a number (A* = 8, A = 7, B = 6 etc. ... G = 1 and U = 0), sum them together and divide by the number of GCSEs.

b There is a difference in the mean GCSE score of children who have siblings when compared with children who have no siblings.

c This would have to be an independent measures design (i.e. different participants in each condition) as it is not possible to be both an only child and have siblings.

d It would be best if all the children had taken their GCSEs in the same year at the same school – so this would control for two extraneous variables (i) overall quality of the school and (ii) standards of GCSE should be more likely to be the same.

e I would approach students in my year at school (Year 12) and ask them to take part in a study on academic achievement. If they said yes, I would ask them to tick the number of A*s, As, Bs etc. they got at GCSE (they do not need to list the subjects) on a slip of paper. Then I would ask them 'Do you have siblings – yes or no?' and note this on the same slip of paper. I would not record the person's name to help keep the data confidential and anonymous. I would then debrief the participant and tell them the aim of the study and see if they had any questions.

2 a I would find a group of students (opportunity sample, having made sure they have not had any chocolate) and show them a video of a popular comedy show, e.g. Vicar of Dibley. I would use a device to measure decibels (sound) which would record how much laughter there is. (More positive mood = more laughter.) I would then, one week later, give them each at least one chocolate bar to eat and half an hour later (to allow time to digest) show them another Vicar of Dibley video (having been previously rated to be of equal funniness as the first episode – could be really clever and counterbalance the episode so half of the participants are shown one episode first and the other participants see the second episode first) and again measure the sound of laughter. This is a repeated measures design because participants take part in both conditions.

b Experimenter bias refers to when the experimenter's expectations of the study translate into cues which then might affect participant behaviour. Therefore, in this study, although the DV (mood) is measured objectively, the experimenter might, when turning on the video or when giving out the chocolate, show that he or she expects that the participants should be in a good mood or find the first video more funny and communicate to the participants by smiling more or being more jolly. Thus the reason for the participants' behaviour is not due to the IV but due to the experimenter's expectations.

c Demand characteristics might affect participants because in the first condition the participants were not given anything at all, whereas in the second condition they were given chocolate. Therefore, the participants might deduce that they should respond more positively and show a better mood during the second video, and behave accordingly.

d You could use a questionnaire (see mood checklist described above) and add a question about how much chocolate the participant typically eats per day. Then it would be possible to e.g. correlate number of chocolate bars consumed per day and score on the mood checklist. Or, divide the respondents into two conditions – those who eat one or less bars per day and those who eat two or more, and see if there is a significant difference between the scores in the two groups.

Qs (page 18)

1 a Students who were identified as 'bloomers' will make greater IQ gains over the course of the school year than students not identified as 'bloomers'.

b This is an independent measures design. This is because there are different participants in each of the two conditions – one set of children were identified as bloomers and another set were not. One strength of using this design is that there could be no 'order effects' from being in one condition and then another, which, in this study would be very difficult to manage or to make believable to participants (i.e. it would be difficult to tell participants that they were bloomers at one point and not bloomers at another point).

c Another possible design would be a matched pairs design. This would mean that each person in the bloomers group would be matched to a person in the ‘normal’ group for IQ score and possibly other things such as gender and social class. This may actually make the research stronger because it then controls for participant variables and means that changes in IQ
Chapter 1 Psychological investigations

Qs (page 20)
1. There will be a significant positive correlation between the hunger and the attractiveness. The range of the attractiveness of food will be from 1 to 10.
2. How hungry are you at this moment? Not at all hungry 1 2 3 4 5 6 7 8 9 10 Extremely hungry
3. There may be differences between people – what is ‘extremely hungry’ for one person might feel the same as just feeling peckish for another person (approx. 5). Also, it is quite a large scale and without defining the numbers in the middle (‘could probably eat something small’/‘feeling a bit peckish’) it is difficult to participants to know how to pick between the numbers in the middle of the scale.
4. From a correlation, you cannot infer cause and effect. This is because a correlation just shows that the two variables co-vary. It is equally as likely (for a correlation) that X causes Y or Y causes X, or that A causes both X and Y. So, from a correlation here, it would be no more valid to conclude that the attractiveness of food causes a person to be hungry than concluding that being hungry causes the food to look better.
5. One ethical issue might be harm. If the food presented is e.g. meat and the person is very big on being e.g. a vegan, they might actually feel offended. Therefore, to avoid this happening, it might be best to ask people before they take part in the study whether there are any types of food they really do not like just to make sure it is not contained within the particular plate of food used in the study.

Qs (page 21)
1. One problem is that a student may feel very motivated, but not like to display it in class e.g. because it looks uncool to answer questions.
2. Another way to assess motivation would be through using a questionnaire and giving statements such as ‘I really want to do well in my A level’ or ‘I am prepared to put in a lot of work in order to get the grade that I want’ with a 5 point scale going from ‘disagree’ to ‘agree’. At the end, this could give a motivation score.
3. Time sampling would be done so that for each participant, every one minute, the observer could note down their activity, e.g. listening, writing, talking to friends, putting up hands, doing work/task set by teacher etc. One strength of this is that it gives an overall sense of the proportion of time spent doing each activity (which just event sampling does not).
4. I would conclude that there is a positive correlation between motivation and average exam scores.

Qs (page 23)
1. This is an example of the scientific method because the researchers wanted to find out if memory might be distorted by a leading question. They designed an experiment to test...
this by comparing the effect of different information (the IV) on people’s responses (the DV). This means that they could be reasonably confident that the behaviour they observed was actually due to the type of question asked. The data from the study is open to scrutiny and the method is clearly explained so that anyone else could try to repeat this experiment.

2 Freud’s study of Little Hans is not replicable. It is a one-off case study of a single boy with his own peculiar set of phobias and circumstances. The way in which Freud talks to him and the discussions of dreams etc. are not at all standardised and so could not be repeated either. Although the boy was observed and data collected from him, there is a strong sense that all that was reported was that which would fit and confirm Freud’s theory – anything which might falsify or challenge the theory is overlooked, underestimated or passed over (e.g. that Hans in seeing a large horse fall may have a behaviourist cause for his phobia, not a psychodynamic one). Therefore, Freud’s explanations for Hans’ behaviour may not have been the best explanations.

3 Thigpen and Cleckley used some aspects of the scientific method, e.g. comparing scores on IQ tests, EEG results etc. between the two personalities. However, it is not scientific because it is still a case study, the data cannot be replicated (it was mostly very personal) and because, whatever doubts one has, it is not possible to falsify their conclusions. Rosenhan uses some aspects of the scientific method, e.g. there are some controls such as all the pseudopatients reported the same auditory hallucinations etc. However, the data collected by the pseudopatients was very anecdotal and unstructured, and participant observations can be unreliable. Also, the whole idea of the method – turning up to the hospital and presenting psychiatric symptoms biases the likelihood of getting the right diagnosis and being admitted to hospital. (It would have been better to show videos of patients describing their symptoms instead and asking doctors to make a diagnosis.)

4 This is frequently debated. Are humans a natural phenomenon? Well, yes. Is the human mind a natural phenomenon? Yes, again. Can we study it in the way we study rocks, chemicals, organisms, animals? Hmmm – to an extent. The problem is that humans, with consciousness, high levels of awareness (including thinking about how to behave and how they should respond) and metacognition (thinking about thinking) means that they are likely to reflexively respond to being studied in a way which a chemical, organism or even an animal might not, i.e. not so ‘natural’. Additionally, humans are in part formed by society – which is always changing. This is not such a big issue for the subject of natural sciences. One other big problem is that of direct observation. On the whole, natural phenomena can be directly observed – chemical reactions, what happens in a cell in the body etc. – but this is not true of the mind. We can only ever guess or infer what is going on in the human mind, even if we use the latest brain scanning we cannot directly see into the activity of the mind. Probably focusing upon the purely behavioural aspects of behaviour is likely to lead to the research being more like a natural science than focusing upon the mind, or reasons behind behaviour.