

# A-Level Psychology Research Methods Exam with A-Grade Answers

By [PsychLogic](#), published Nov 12, 2019

---

Read the item and then answer the questions that follow.

Following previous research indicating the social benefits of green space in urban areas, two psychology students decided to observe social behaviour in public spaces. They focused on two neighbouring towns, Greensville where most public spaces were planted with flowers and vegetables, and Brownton where most public spaces were paved with concrete.

The students compared the instances of considerate behaviours in the two towns.

Considerate behaviour categories included putting litter in the bin, having a dog on a lead and riding a bike with care.

The observations were carried out in four different areas of a similar size in each town on weekdays between the hours of 4.30pm and 6.00pm. The students worked together to ensure inter-observer reliability, recording each target behaviour whenever it occurred.

(48 marks)

- (a) Should the hypothesis for this study be directional? Explain your answer.

(2 marks)

The hypothesis should be directional as previous research indicates social benefits of green space.

- (b) Before the observation could begin, the students needed to operationalise the behaviour category 'riding a bike with care'. Explain what is meant by operationalisation and suggest two ways in which 'riding a bike with care' could have been operationalised.

(4 marks)

Operationalisation means finding a way of measuring a variable in a numerical/quantitative form. The researchers could count the number of time people ride a bike at 20mph or less or how many people ride on the correct side of the road only.

- (c) The students thought that having a dog on a lead was a useful measure of considerate behaviour because it had face validity. Explain what is meant by face validity in this context. (3 marks)

Face validity means whether the form of measurement we are using appears to be an accurate form of measurement. Having a dog on a lead appears to be measuring considerate behaviour because if a dog is on a lead it is less able/likely to upset other people by frightening, chasing, biting, etc.

- (d) Identify and briefly outline two other types of validity in psychological research.

(4 marks)

Content Validity involves independent experts being asked to assess the validity/accuracy of instruments/tests used to measure a variable: e.g. agreeing that a particular IQ test is a valid measure of intelligence.

Concurrent Validity involves comparing the validity of a new test/measure against an established test/measure whose validity is already known and trusted. For example, the results of a new form of IQ test could be tested against an old, established IQ test. If scores correlate between the 2 tests they are said to have concurrent validity.

- (e) Identify the behaviour sampling method used by the students. Shade one box only.

- A. Time sampling
- B. Pair sampling
- C. Event sampling
- D. Target sampling

(1 mark)

C

(The reason is because *they count every time a behaviour/event occurs: e.g. putting litter in the bin, having a dog on a lead, riding a bike with care*).

- (f) Explain how inter-observer reliability could be ensured by working as a pair.

(3 marks)

We can try to ensure that the 2 observers are defining behaviours and recording observations in the same way as each other. Thus, before the study begins observers should be trained through the use of, for example, a training video where they learn and are then tested on how to define and categorise behaviours in the same way as each other.

We can assess inter-rater reliability by analysing the correlation between different observers score when measuring the same behaviour. This will produce a correlation coefficient (see Correlation Studies and Spearman's rho test): e.g. +0.9 = a strong positive correlation (they are rating things in the same way as each other).

The data for considerate behaviours is shown in Table 1.

Table 1

Considerate behaviours			
	Litter in bin	Dog on lead	Riding bike with care
Greensville	23	12	19
Brownton	10	17	9

The students noted that overall more considerate behaviours occurred in Greensville than in Brownton.

- (g) Calculate the ratio of considerate behaviours observed in Greensville to considerate behaviours observed in Brownton. Show your workings and present your answer in the simplest form.

(3 marks)

Add up all considerate behaviours for both towns. Greensville =  $23 + 12 + 19 = 54$ ; Brownton =  $10 + 17 + 9 = 36$ . Then simplify  $54/36$ . They are both divisible by 9.  $54/9 = 6$ ;  $36/9 = 4$ .  $6/4$  can then be simplified to the ratio of 3:2.

- (h) The students carried out a Chi-square test on their data. Explain why the Chi-square test was an appropriate test to use in this case

(3 marks)

Because they are testing for a difference using an independent groups design with nominal data.

(Data is nominal because they are counting categories of behaviour (e.g. putting litter in a bin) not giving individuals a score for a behaviour.)

- (i) In order to interpret the results of the Chi-square test the students first needed to work out the degrees of freedom. They used the following formula.

$$\text{Degrees of freedom (df)} = (r-1) \times (c-1)$$

$r$  = number of rows and  $c$  = number of columns

Calculate the degrees of freedom for the data in Table 1. Show your workings.

(2 marks)

AQA do not require you to know what ‘degrees of freedom are’ and this question does not ask you to explain what they are, it simply asks you to make a simple calculation.

Answer.

“For table 1:

There are 2 rows – therefore  $r = 2$

There are 3 columns – therefore  $c = 3$

Therefore,  $(df) = (2-1) \times (3-1)$

Therefore,  $(df) = 1 \times 2$

Therefore  $(df) = 2$ ”

- (j) The calculated value of Chi-square was 6.20. Referring to Table 2 below, state whether or not the result of the Chi-square test is significant at the 0.05 level of significance. Justify your answer.

(3 marks)

Table 2

df	Levels of significance for a one-tailed test			
	0.10	0.05	0.025	0.01
	Levels of significance for a two-tailed test			
	0.20	0.10	0.05	0.02
1	1.64	2.71	3.84	5.41
2	3.22	4.60	5.99	7.82
3	4.64	6.25	7.82	9.84
4	5.99	7.78	9.49	11.67

To be significant at the level shown the calculated value of Chi Square must be equal to or greater than the critical/table value.

*Answer.*

*“The results are significant because the calculated value of Chi-square (6.20) > the CV (4.60) (1-tailed test,  $p = 0.05$ ,  $df = 2$ ). Therefore, the experimental hypothesis should be accepted.”*

Explanation + notes.

The above is the quickest way of writing the correct answer. The way to work out the correct answer is:

1. What is the observed/calculated value? In this case it is the value 6.20.
2. Find the correct critical value (CV) on the table. You need to cross-reference the correct level of significance (we are told it is 0.05) for the correct one/two-tailed test (we are told it's one-tailed) with the correct df value (which we have previously calculated as 2). Cross-referencing this gives us the CV value of 4.60.
3. The rule for the Chi-square test (written under the table) is that if the Chi-square value > the CV then the result is significant. 6.40 is greater than the CV (6.40 > 4.60) therefore the result is significant.

- (k) In the discussion section of their report of the investigation the students wanted to further discuss their results in relation to levels of significance.

Write a short paragraph the students could use to do this.

(4 marks)

Although the question mentions the discussion section, you are asked to discuss results in relation to levels of significance only – so only write about this.

Answer.

“Results are significant at  $p=0.05$ . This means the researchers can be 95% confident that the difference found between considerate behaviours in the 2 towns is a real/true difference. However, there is a 5% probability that these results occurred due to chance and, therefore, a 5% chance that the researchers made a type 1 error – accepted the experimental hypothesis when, in fact, these results simply occurred due to chance and there is no real difference between the 2 towns.

Results are also significant at the more stringent  $p=0.025$  level of significance (Chi-square value of  $6.40 > CV$  of  $5.99$ ) therefore the researchers can be 97.5% confident that there is a significant difference. However, results are not significant at  $p=0.01$  (Chi-square value of  $6.40 < CV$  of  $7.82$ ).”

- (1) Design an experiment to investigate the effect of indoor plants on mood in office workers. For your measure of mood, you should devise a measure that would give data suitable for testing at the ordinal level of measurement.

In your answer you should provide details of:

Design – include reference to the experimental design, variables and controls

Materials/Apparatus – describe any special materials required

Data analysis that could be used – include reference to descriptive and inferential analysis.

Justify your choices

(12 marks)



General guidance – it must be an experiment, data produced must be at the ordinal level, you must cover all 3 bullet-points... For 12 marks, you should write for at least 15 minutes so they expect lots of depth and detail but do not go off track: for example, you are not asked to give details of sampling techniques... Of particular importance is that there must be sufficient detail for the study to be replicated by someone reading your answer.

**Answer.**

“This experiment would use an independent groups design where 2 groups of 30 office workers in 2 separate buildings would be compared. The IV would be the presence of indoor plants in 1 of the offices and the absence of plants in the other office. The DV would be office workers mood. Offices would be chosen which were open-plan and contained 30 workers each. The two groups of workers would work in the same occupation: for example, journalism (to eliminate differences in job as a potential extraneous variable) and workers in the office chosen to contain plants would be pre-tested to ensure none of the workers were averse to or allergic to plants (to eliminate this as an extraneous variable). Anyone who was averse to or allergic to plants would be removed from the study. The DV would be operationalised by constructing a closed-ended questionnaire composed of 20 questions asking participants about mood with 5 answers for each question ranging from 1 (strongly disagree) to 5 (strongly agree). Questions would be phrased so that ‘strongly agree’ answers indicated a positive mood. An example of a question is: ‘When you wake up in the morning do you generally feel happy and content?’ This questionnaire would be given to all office workers at the start of the study so that they all received a mood score of  $x/100$ .

The study would continue for 2 months. In the ‘plant’ condition, the office workers’ office would be filled with a variety of plants. These would be well-cared for by someone other than the office workers. This would eliminate dead/ugly plants as a potential extraneous variable. After 2 months, all office workers would take the mood questionnaire again and the positive or negative difference in each individual’s mood score would be calculated. This would produce ordinal data: for example, office worker number 1 in the plant condition started with a mood score of  $50/100$  and ended with a mood score of  $65/100$ , therefore, their mood increase/decrease score is +15.

Data could be analysed in various ways. The mean mood score for both groups could be calculated and compared along with a range and SD. Data could also be analysed using the Mann-Whitney U test as the study is a test of difference using an independent groups design with ordinal data.”