



# Psychology

**AS Level**

**Psychology**

Unit G542

Exemplar Candidate Work  
from June 2014

May 2015

# CONTENTS

<b>SECTION A – QUESTION 1</b>	8
SAMPLE ANSWER(S)	8
COMMENTARY	8
<b>SECTION A – QUESTION 2(a)</b>	10
SAMPLE ANSWER(S)	10
COMMENTARY	10
<b>SECTION A – QUESTION 2(b)</b>	11
SAMPLE ANSWER(S)	11
COMMENTARY	11
<b>SECTION A – QUESTION 3</b>	12
SAMPLE ANSWER(S)	12
COMMENTARY	12
<b>SECTION A – QUESTION 4</b>	14
SAMPLE ANSWER(S)	14
COMMENTARY	14
<b>SECTION A – QUESTION 5</b>	16
SAMPLE ANSWER(S)	16
COMMENTARY	16
<b>SECTION A – QUESTION 6(a)</b>	18
SAMPLE ANSWER(S)	18
COMMENTARY	18
<b>SECTION A – QUESTION 6(b)</b>	19
SAMPLE ANSWER(S)	19
COMMENTARY	19
<b>SECTION A – QUESTION 7(a)</b>	20
SAMPLE ANSWER(S)	20
COMMENTARY	20
<b>SECTION A – QUESTION 7(b)</b>	22
SAMPLE ANSWER(S)	22
COMMENTARY	22
<b>SECTION A – QUESTION 8</b>	24
SAMPLE ANSWER(S)	24
COMMENTARY	24

<b>SECTION A – QUESTION 9</b>	26
SAMPLE ANSWER(S)	26
COMMENTARY	26
<b>SECTION A – QUESTION 10(a)</b>	28
SAMPLE ANSWER(S)	28
COMMENTARY	28
<b>SECTION A – QUESTION 10(b)</b>	29
SAMPLE ANSWER(S)	29
COMMENTARY	29
<b>SECTION A – QUESTION 11</b>	31
SAMPLE ANSWER(S)	31
COMMENTARY	31
<b>SECTION A – QUESTION 12</b>	33
SAMPLE ANSWER(S)	33
COMMENTARY	33
<b>SECTION A – QUESTION 13</b>	35
SAMPLE ANSWER(S)	35
COMMENTARY	35
<b>SECTION A – QUESTION 14</b>	37
SAMPLE ANSWER(S)	37
COMMENTARY	37
<b>SECTION A – QUESTION 15</b>	39
SAMPLE ANSWER(S)	39
COMMENTARY	39
<b>SECTION B – QUESTION 16(a) IN RELATION TO DEMENT &amp; KLEITMAN</b>	41
SAMPLE ANSWER(S)	41
COMMENTARY	41
<b>SECTION B – QUESTION 16(b) IN RELATION TO DEMENT &amp; KLEITMAN</b>	43
SAMPLE ANSWER(S)	43
COMMENTARY	43
<b>SECTION B – QUESTION 16(c) IN RELATION TO DEMENT &amp; KLEITMAN</b>	45
SAMPLE ANSWER(S)	45
COMMENTARY	46

<b>SECTION B – QUESTION 16(d)</b>	IN RELATION TO DEMENT & KLEITMAN	47
SAMPLE ANSWER(S)		47
COMMENTARY		48
<b>SECTION B – QUESTION 16(e)</b>	IN RELATION TO DEMENT & KLEITMAN	49
SAMPLE ANSWER(S)		49
COMMENTARY		50
<b>SECTION B – QUESTION 16(f)</b>	IN RELATION TO DEMENT & KLEITMAN	51
SAMPLE ANSWER(S)		51
COMMENTARY		52
<b>SECTION B – QUESTION 16(a)</b>	IN RELATION TO THIGPEN AND CLECKLEY	53
SAMPLE ANSWER(S)		53
COMMENTARY		53
<b>SECTION B – QUESTION 16(b)</b>	IN RELATION TO THIGPEN AND CLECKLEY	54
SAMPLE ANSWER(S)		54
COMMENTARY		54
<b>SECTION B – QUESTION 16(c)</b>	IN RELATION TO THIGPEN AND CLECKLEY	56
SAMPLE ANSWER(S)		56
COMMENTARY		57
<b>SECTION B – QUESTION 16(d)</b>	IN RELATION TO THIGPEN AND CLECKLEY	58
SAMPLE ANSWER(S)		58
COMMENTARY		59
<b>SECTION B – QUESTION 16(e)</b>	IN RELATION TO THIGPEN AND CLECKLEY	60
SAMPLE ANSWER(S)		60
COMMENTARY		61
<b>SECTION B – QUESTION 16(f)</b>	IN RELATION TO THIGPEN AND CLECKLEY	62
SAMPLE ANSWER(S)		62
COMMENTARY		63
<b>SECTION B – QUESTION 16(a)</b>	IN RELATION TO REICHER AND HASLAM	64
SAMPLE ANSWER(S)		64
COMMENTARY		64
<b>SECTION B – QUESTION 16(b)</b>	IN RELATION TO REICHER AND HASLAM	65
SAMPLE ANSWER(S)		65
COMMENTARY		65

<b>SECTION B – QUESTION 16(c) IN RELATION TO REICHER AND HASLAM</b>	67
SAMPLE ANSWER(S)	67
COMMENTARY	68
<b>SECTION B – QUESTION 16(d) IN RELATION TO REICHER AND HASLAM</b>	69
SAMPLE ANSWER(S)	69
COMMENTARY	70
<b>SECTION B – QUESTION 16(e) IN RELATION TO REICHER AND HASLAM</b>	72
SAMPLE ANSWER(S)	72
COMMENTARY	72
<b>SECTION B – QUESTION 16(f) IN RELATION TO REICHER AND HASLAM</b>	73
SAMPLE ANSWER(S)	73
COMMENTARY	74
<b>SECTION C – QUESTION 17(a) IN RELATION TO THE BEHAVIOURIST PERSPECTIVE</b>	75
SAMPLE ANSWER(S)	75
COMMENTARY	75
<b>SECTION C – QUESTION 17(b) IN RELATION TO THE BEHAVIOURIST PERSPECTIVE</b>	76
SAMPLE ANSWER(S)	76
COMMENTARY	76
<b>SECTION C – QUESTION 17(c) IN RELATION TO THE BEHAVIOURIST PERSPECTIVE</b>	77
SAMPLE ANSWER(S)	77
COMMENTARY	78
<b>SECTION C – QUESTION 17(d) IN RELATION TO THE BEHAVIOURIST PERSPECTIVE</b>	80
SAMPLE ANSWER(S)	80
COMMENTARY	81
<b>SECTION C – QUESTION 18(a) IN RELATION TO THE COGNITIVE APPROACH</b>	83
SAMPLE ANSWER(S)	83
COMMENTARY	83
<b>SECTION C – QUESTION 18(b) IN RELATION TO THE COGNITIVE APPROACH</b>	84
SAMPLE ANSWER(S)	84
COMMENTARY	84
<b>SECTION C – QUESTION 18(c) IN RELATION TO THE COGNITIVE APPROACH</b>	86
SAMPLE ANSWER(S)	86
COMMENTARY	87

<b>SECTION C – QUESTION 18(d) IN RELATION TO THE COGNITIVE APPROACH</b>	88
SAMPLE ANSWER(S)	88
COMMENTARY	89

The exemplars in this document have been re-typed for ease of reading but are taken from real candidate work and therefore include the original spelling and grammatical errors made by the candidates at the time.



We'd like to know your view on the resources we produce. By clicking on the 'Like' or 'Dislike' button you can help us to ensure that our resources work for you. When the email template pops up please add additional comments if you wish and then just click 'Send'. Thank you.

If you do not currently offer this OCR qualification but would like to do so, please complete the Expression of Interest Form which can be found here: [www.ocr.org.uk/expression-of-interest](http://www.ocr.org.uk/expression-of-interest)

**OCR Resources: the small print**

OCR's resources are provided to support the teaching of OCR specifications, but in no way constitute an endorsed teaching method that is required by the Board and the decision to use them lies with the individual teacher. Whilst every effort is made to ensure the accuracy of the content, OCR cannot be held responsible for any errors or omissions within these resources.

© OCR 2015 - This resource may be freely copied and distributed, as long as the OCR logo and this message remain intact and OCR is acknowledged as the originator of this work.

OCR acknowledges the use of the following content:

Thumbs up: alex\_white/Shutterstock.com, Thumbs down: alex\_white/Shutterstock.com

Please get in touch if you want to discuss the accessibility of resources we offer to support delivery of our qualifications: [resources.feedback@ocr.org.uk](mailto:resources.feedback@ocr.org.uk)

# INTRODUCTION

## General feedback on performance

Examiners considered the paper to be fair and accessible with the standard being in line with previous sessions. Overall, there was an acceptable range of marks across both candidates and the paper with good candidates showing an excellent understanding of both psychology and the core studies.

There was little evidence of candidates running out of time and there were very few rubric errors.

All sections of this paper assess a candidate's understanding of key psychological terms and concepts and require this understanding to be demonstrated by supporting answers with appropriate evidence from either the core studies themselves and/or the background to the core studies/theories on which the studies are based. There were many instances that suggested candidates did not know or understand the theories and research surrounding the core studies and/or adequate specific details of these studies.

Unfortunately, many candidates lost valuable marks in both questions 17/18 c and 17/18 d by supporting their identified similarities/differences and strengths/weaknesses with evidence that was not appropriate. Many marks were also lost in question 17/18 d by candidates failing to respond to the requirements of the question and instead of discussing strengths/weakness of laboratory experiments, discussing strengths/weaknesses of either the behaviourist perspective or the cognitive approach.

The quality of written communication and writing, as in previous sessions, prevented some candidates from attaining higher marks. As suggested in previous reports, such candidates may be eligible for access arrangements. Furthermore, there were more instances than in previous sessions of candidates writing below the given lines, sideways in margins, and in additional booklets when there was sufficient space on the extra pages for their extended answers. Candidates should be encouraged to write only on the lines provided and to use the additional pages at the back of the script before moving on to additional booklets.

# SECTION A, QUESTION 1

## Question 1

From Baron-Cohen, Jolliffe, Mortimore and Robertson's study on autism in adults:

Describe the Basic Emotion Recognition Task used in this study.

[4 marks]

## SAMPLE ANSWERS

(a) The participants were shown pictures of whole faces of people displaying the six basic emotions e.g. happy, sad, angry. They were all shown the same pictures and asked to say how the participant was feeling. They could choose from two words under the picture to best describe how the person was feeling. The words were opposite to each other e.g. happy / sad.

**Mark – 4 marks**

(b) The Basic Emotion Recognition Task contained a set of standard photos from magazines. Participants were asked which of the six basic emotions (e.g. happy, sad, disgust, anger) the face was showing. It was used to test the emotion perception ability of the autistic adults

**Mark – 3 marks**

(c) The participants were asked to identify the emotions of whole faces. The basic emotions included disgust and fear.

**Mark – 2 marks**

(d) Participants were shown eyes of people. The picture showed both eyes, the top of the nose and the brow. Each picture showed the eyes of a basic emotion e.g. happy, sad, surprised, and participants were asked to identify the emotion shown.

**Mark – 0 marks**

## COMMENTARY

(a) The full 4 marks were awarded here because it is an accurate description which refers to: the pictures being of whole faces, the six basic emotions, appropriate examples of the emotions tested and the forced choice option available when having to identify the emotion shown.

(b) 3 marks were awarded here because this is an accurate description, but as the candidate has not referred to the forced choice option when having to identify the emotion shown, full marks could not be gained.

(c) This response was awarded 2 marks. The candidate has stated that the pictures used for this test were of the whole face and given two appropriate examples of emotions used. However, as the candidate has not referred to either the forced choice option when having to identify the emotion shown or that there were six basic emotions/that the photos were based on the Ekman categories, the answer was considered partial and awarded 2 marks.

(d) 0 marks were awarded here as the candidate has referred to pictures of eyes and not the face.

### **How the answer could be improved**

The Basic Emotion Recognition Task involved judging photographs of whole faces displaying basic emotions (based on the Ekman categories). Six faces were used testing the following basic emotions: happy, sad, angry, afraid, disgusted, and surprised. Participants had to state, from a forced choice of two mental states, which emotion was shown in each photograph.

### **General performance on the question**

Many candidates scored well on this question though few included reference to the participants having to state, from a forced choice of two mental states, which emotion was shown on each photograph. There were also many instances where candidates either described the Eyes task or gave a description which could have applied to either the Basic Emotions Task or the Eyes Task because they failed to state that the photographs were of whole faces.

## SECTION A, QUESTION 2 (a)

### Question 2 (a)

From Savage-Rumbaugh's study into symbol acquisition by pygmy chimpanzees:

Outline how data was recorded in this study whilst the pygmy chimpanzees were indoors.

[2 marks]

## SAMPLE ANSWERS

(a) When the chimpanzees were indoors, data was recorded using the electronic lexigram board which was wired to a computer which would automatically record the pygmy chimpanzees' utterances.

**Mark – 2 marks**

(b) Data was recorded from the lexigram.

**Mark – 1 mark**

(c) Observation was used to record data whilst the pygmy chimps were indoors. Keepers may ask the chimps a question and utterances were recorded into various classifications.

**Mark – 0 marks**

## COMMENTARY

(a) The full 2 marks were awarded as the candidate has referred to the use of the lexigram and that as this was linked to a computer, data would be automatically recorded.

(b) 1 mark was awarded here for a vague/partial answer. The candidate has merely identified how data was gathered. This is not an outline and so needed further development / explanation.

(c) 0 marks were awarded as this was not the way data was gathered whilst the pygmy chimps were indoors.

### How the answer could be improved

When indoors all lexigrams used by Kanzi and Mulika were automatically recorded by computer-monitored keyboards.

### General performance on the question

This question was generally well answered though some candidates referred to data being recorded through observation of lexigram usage; this was not creditworthy as this was how data was gathered whilst outdoors.

## SECTION A, QUESTION 2 (b)

### Question 2 (b)

From Savage-Rumbaugh's study into symbol acquisition by pygmy chimpanzees:

Suggest one strength of the method used to gather data whilst the pygmy chimpanzees were indoors.

[2 marks]

## SAMPLE ANSWERS

(a) The data collected was objective as it was exactly which symbols the pygmy chimpanzees had pressed. This could not be affected by human error.

**Mark – 2 marks**

(b) One strength of this method is that it allowed for quantitative data to be collected which can be statistically analysed.

**Mark – 1 mark**

(c) The observational method used gave the study rich qualitative (responses) and quantitative (number of responses which were appropriate) data.

**Mark – 0 marks**

## COMMENTARY

(a) The full 2 marks were awarded here as an appropriate strength has been identified (an objective measure which will not be susceptible to human error) and clearly described within the context of the study.

(b) Only 1 mark could be awarded here because, although an appropriate strength has been identified (quantitative data could be gathered which is easy to analyse), the answer has not been contextualised so only partial marks could be awarded to the named study.

(c) 0 marks were awarded here. This response is not creditworthy as the candidate is referring to observation.

### How the answer could be improved

The use of an electronic, computerised lexigram to record symbol usage is an objective measure which removes any researcher bias or chances of human error in the recording of data.

### General performance on the question

This question was again, generally well answered with many candidates giving clear, fully contextualised responses.

## SECTION A, QUESTION 3

### Question 3

In their study on eyewitness testimony, Loftus and Palmer proposed that two kinds of information form one's memory of an event.

Describe these two kinds of information in relation to this study.

[4 marks]

### SAMPLE ANSWERS

(a) One type of information that Loftus and Palmer proposed is the initial perception of the event. This is what the person sees themselves at the time of the event. In this study, for example, it was the viewing of the seven film clips of car accidents shown in Experiment 1. The other type of information that Loftus and Palmer proposed is the external information received after the event. For example in Loftus and Palmer's study this information would be the inclusion of glass in the crash from the question, "Did you see any broken glass?" when participants returned a week later to complete a questionnaire in Experiment 2.

**Mark – 4 marks**

(b) One type of information is the memory of the event at the time. In this case the car crashes the participants saw on the slide show. The second type of information is the post event information such as the leading questions.

**Mark – 3 marks**

(c) The memory of the event when it was witnessed – information at the time – and the memory of the event after a period of time from being witnessed – information given after the event.

**Mark – 2 marks**

(d) If there was any external object e.g. broken glass. If someone asks a leading question or uses a strong/weak verb e.g. smashed.

**Mark – 0 marks**

### COMMENTARY

(a) The full 4 marks were awarded here. It is a clear, fully contextualised description of both types of information identified by Loftus and Palmer.

(b) 3 marks were awarded. One type of information (memory gained at the time of the event) has been identified and described in the context of Loftus and Palmer's study. However, although the second type of information (post event information) has been identified, it has not been described adequately in context of the named study i.e. the reference to 'leading questions' is too vague and needed further elaboration/development.

(c) Although both kinds of information have been identified, neither has been described in the context of Loftus and Palmer's study; only 2 marks could be awarded as this is an uncontextualised answer.

(d) 0 marks were awarded here. Although the candidate has referred to 'leading questions', this has not been linked to the possibility that such a question may lead to memory distortion. The response was therefore deemed not creditworthy.

### **How the answer could be improved**

Memory is determined by two sources:

- (i) one's own perception gleaned at the time of the original event i.e. in Loftus and Palmer's study, whilst watching the film(s) of the car accidents
- (ii) external information supplied after the fact i.e. in this study, through the leading questions in relation to vehicular speed and/or the recollection of seeing broken glass.

### **General performance on the question**

A generally well answered question. However some candidates referred to 'schemas' which was not one of the types of information proposed by Loftus and Palmer in this study.

Furthermore, many candidates only made reference to the use of leading questions which was an example of the second type of information that forms an individual's memory of an event – information received after the event.

# SECTION A, QUESTION 4

## Question 4

From Bandura, Ross and Ross' study of aggression:

Describe how individual differences in levels of aggression were controlled for in this study.

[4 marks]

## SAMPLE ANSWERS

(a) To control individual differences in levels of aggression, each child was rated on a 5-point scale before the study on their levels of aggression by their school/nursery teacher and one researcher. They were then all put into groups of three matched on their aggression level and one from each group was allocated to either one of the model conditions – aggressive, non-aggressive - or the control condition where no model was observed.

**Mark – 4 marks**

(b) The children were put into a group of 6 based on their pre-existing levels of aggression. They were observed in the school playground by their teacher and one of the researchers who rated them out of 5 based on their verbal and physical aggression. The scores of the different children were collected and then children were matched with other children with similar aggression scores. This controlled for individual differences in levels of aggression.

**Mark – 3 marks**

(c) The children were pre-rated on for aggression and then matched on their levels of aggression.

**Mark – 2 marks**

((d) The levels of aggression were controlled by having the same toys for each condition so the children could react to the same toys as the model did in similar ways.

**Mark – 0 marks**

## COMMENTARY

(a) This is an accurate, fully contextualised description and was awarded the full 4 marks. Reference has been made to: the pre-rating of the children's aggression levels using a 5-point scale; the people who did the rating; that children were then matched on their assessed levels of aggression and that they were then evenly distributed between the three conditions. Full marks were therefore awarded.

(b) Unfortunately this candidate has failed to state that once the children had been matched on the assessed levels of aggression, they were then evenly/randomly distributed between the three conditions. The response was therefore awarded 3 marks.

(c) 2 marks were awarded here. The candidate has stated that the children were pre-rated for aggressive tendencies but has failed to say how or by whom this was done. They have also failed to say how the children were distributed amongst the different conditions once they had been matched on aggressiveness.

(d) 0 marks were awarded. This is an incorrect answer as this answer refers to the children's aggression levels being aroused.

### **How the answer could be improved**

Participants in the experimental and control groups were matched individually on the basis of ratings of their aggressive behaviour in social interactions in the nursery school. They were rated on four 5-point rating scales by the experimenter and a nursery school teacher, both of whom were well acquainted with the children. These scales measured the extent to which the children displayed physical aggression, verbal aggression, aggression towards inanimate objects, and aggressive inhibition. (The latter scale, which dealt with the child's tendency to inhibit aggressive reactions in the face of high instigation, provided a measure of aggression anxiety.)

51 children were rated independently by both judges so as to permit an assessment of inter-rater agreement). The composite score was obtained by summing the ratings of the four aggression scales and on the basis of these scores the children were arranged in triplets and then randomly assigned to either one of the two experimental conditions – aggressive or non-aggressive - or the control condition.

### **General performance on the question**

Many candidates scored well on this question by referring to (a) by whom and where the children were observed and rated for aggression (b) how they were rated for aggressive tendencies (c) the fact that they were then matched on aggression levels. Few candidates however said that the children were then arranged in triplets and randomly assigned to either one of the two experimental conditions or the control condition.

## SECTION A, QUESTION 5

### Question 5

From Samuel and Bryant's study into conservation:

Describe two ways in which the results of this study support Piaget's theory of cognitive development.

[4 marks]

### SAMPLE ANSWERS

(a) Samuel and Bryant's study shows that children's ability to conserve improves as they get older. For example, 8-year olds made fewer errors than 5-year-olds on all the conservation tasks. This supports Piaget's theory which says that conservation improves with age. Samuel and Bryant's results also show that children aged between 5 and 8 found the conservation of number easier than that of either mass or volume. This also supports Piaget's theory which says that conservation develops in stages.

**Mark** – 4 marks

(b) Samuel and Bryant's results show that the child's ability to conserve develops with age. This was shown by fewer errors being made by the older children in all of the conservation tasks. The easiest task with the least number of mean errors was the number task.

**Mark** – 3 marks

(c) The older children performed the best in all tasks proving the conservation skills in children improve with age. Another way it supports Piaget's theory is that the children performed the worst in the standard tasks as they had been asked the same question twice.

**Mark** – 2 marks

(d) Children in Samuel and Bryant's study performed better on every task as they got older. Children made the least number of errors on the standard question.

**Mark** – 1 mark

### COMMENTARY

(a) Two appropriate results have been given with both being appropriately linked to Piaget's theory. The response was therefore awarded 2+2=4 marks.

(b) One appropriate result, clearly linked to Piaget's theory has been described. The reference to the number task being the easiest is a mere finding with no link to Piaget's theory. The response was therefore awarded 2+1=3 marks.

(c) The second suggestion does not think to Piaget's theory – it refers to Piaget's methodology. Therefore there is only one creditworthy result which has been appropriately linked to Piaget's theory. The response was therefore awarded 2 marks.

(d) The first suggestion is a mere result from Samuel and Bryant's study, it has not been linked to Piaget's theory, so is only a partial answer, gaining 1 mark. The second suggestion refers to Piaget's methodology which is not creditworthy.

### **How the answer could be improved**

Piaget's theory proposed that the ability to conserve increases with age. Findings from this study support this as the older children (8 year-olds) performed better on all conservation tasks than the younger children (5 year-olds). Piaget's theory also proposed that the ability to conserve develops in stages / differs according to the difficulty of the task. Findings from this study support this as the mean number of errors in all three conditions - one question, standard two question and fixed-array - was lower in the numbers task than in either the mass or volume task

### **General performance on the question**

This question was not well answered. Many candidates either failed to link findings from Samuel and Bryant's study to Piaget's theory or just gave findings from Samuel and Bryant's study and repeated the strap-line – these results support Piaget's theory. Some candidates also failed to describe two ways in which the results of this study supported Piaget's theory of cognitive development.

## SECTION A, QUESTION 6(a)

### Question 6(a)

From Freud's study of Little Hans:

Outline Hans' giraffe fantasy.

[2 marks]

## SAMPLE ANSWERS

(a) There was a tall giraffe and a crumpled giraffe. Little Hans sat on the crumpled giraffe and the tall giraffe called out.

**Mark** – 2 marks

(b) Hans saw 2 giraffes, a tall one and a small, crumpled one. The tall one was sitting on the little one crying out.

**Mark** – 1 marks

(c) Hans was scared that a giraffe would steal something of his.

**Mark** – 0 marks

## COMMENTARY

(a) This is a clear and adequately accurate description of the giraffe fantasy. The full 2 marks were awarded.

(b) This answer is only partially correct. The tall giraffe was not sitting on the little giraffe, Hans sat on the little/crumpled giraffe which caused the big giraffe to cry out. 1 mark was therefore awarded.

(c) 0 marks were awarded here as this is an incorrect answer.

### How the answer could be improved

'In the night there was a big giraffe in the room and a crumpled one, and the big one called out because I took the crumpled one away from it. Then it stopped calling out, and then I sat down on top of the crumpled one.'

### General performance on the question

This was a generally well answered question though some candidates referred to there being three giraffes – Hans' father, mother and himself.

## SECTION A, QUESTION 6(b)

### Question 6(b)

From Freud's study of Little Hans:

Outline Freud's explanation for Hans' giraffe fantasy.

[2 marks]

### SAMPLE ANSWERS

(a) It is the re-working of the morning routine in Hans' parents' bed. Hans liked getting into their bed but his father always objected to this. The big giraffe was his father and the crumpled giraffe was his mother. The big giraffe was calling out because Hans was taking the crumpled giraffe – his mother – away from his father because he wanted her to himself.

**Mark – 2 marks**

(b) Hans associated the big giraffe with his father and wanted to keep him away from his mother.

**Mark – 1 marks**

(c) Hans felt threatened by his father, who is the giraffe with the long neck. This shows the boy's castration phobia.

**Mark – 0 marks**

### COMMENTARY

(a) The full 2 marks were awarded as this is a clear, fully contextualised explanation of the giraffe fantasy.

(b) 1 mark was awarded for a partial answer. The candidate has not explained that the crumpled giraffe was supposed to represent his mother whom he wished to have for himself.

(c) This is an incorrect response. 0 marks were awarded.

#### How the answer could be improved

Freud believed it was a replication of what happened in the morning when Hans tried to get into bed with his parents and suggested this was Hans' distorted way of communicating his wishes in regard to his mother. The big giraffe represented his father whilst the crumpled giraffe represented his mother. By taking the crumpled giraffe away from the big giraffe Hans was showing his wish to have his mother all to himself.

#### General performance on the question

As always with any Freud question, there were many imaginative answers! Many candidates scored very well on this question but there were many who seemed to think that the big giraffe's neck represented Hans' father's penis and the crumpled giraffe his mother's vagina. This was not how Freud, at least in this study, interpreted this fantasy. Some candidates who used the Oedipus complex as an explanation failed to make any link with the giraffe fantasy.

## SECTION A, QUESTION 7(a)

### Question 7(a)

From Sperry's 'split brain' study into the psychological effects of hemisphere disconnection:

Describe one finding from the visual tests used in this study.

[2 marks]

### SAMPLE ANSWERS

(a) Sperry found that when an image was presented to the participant's left visual field the participant could not name the object. He also found that the participant would often say they saw a 'flash of light' when this stimulus was used.

**Mark – 2 marks**

(b) When the participants were shown a picture of a nude woman exposed to the right hemisphere, they laughed and blushed and were unable to say what they had seen.

**Mark – 1 mark**

(c) When an image was flashed to the left visual field the left hemisphere would be able to describe it but the right would act as if it was an unfamiliar vision.

**Mark – 0 marks**

### COMMENTARY

(a) This was awarded the full 2 marks. It is a clear and accurate description of one of Sperry's findings from the visual tests.

(b) 1 mark was awarded here for a partial answer. The candidate needed to include that to get this result information had to be presented to the left visual field.

(c) This is a muddled and incorrect answer. Information presented to the left visual field would be processed by the right hemisphere and a participant was able to name the object shown. The reference to the right hemisphere is unclear. No understanding has been shown in this answer so 0 marks were awarded.

**How the answer could be improved**

- Visual material presented to the RVF is processed by the left hemisphere so can be described in speech and/or writing/ drawing/ pointing using the right hand.
- If an image/word has been identified and responded to in one visual field it can only be recognised again if it is presented to the same visual field.
- Visual material presented to the LVF cannot be identified in speech and/or pointing/ drawing/ writing with the right hand but can be identified by pointing/ drawing/ writing with the left hand.

**General performance on the question**

A well answered question, with only a few candidates referring to 'eye' instead of 'visual field'.

## SECTION A, QUESTION 7(b)

### Question 7(b)

From Sperry's 'split brain' study into the psychological effects of hemisphere disconnection:

Describe one finding from the tactile tests used in this study.

[2 marks]

### SAMPLE ANSWERS

(a) One finding from the tactile tests was that if two different objects were placed in a participant's left and right hands, then the participant's hands could only find their respective object in a grab bag, and not the object that the other hand held.

**Mark** – 2 marks

(b) The participant could only find the object with the same hand.

**Mark** – 1 mark

(c) When an object was placed in the left hand the left hemisphere was able to name it.

**Mark** – 0 marks

### COMMENTARY

(a) This is a clear and accurate description of one of Sperry's findings from the tactile tests. The full 2 marks were therefore awarded.

(b) This is a vague answer, but not incorrect, so 1 mark was awarded. The candidate needed to either say 'if an object was placed in one hand, it could then subsequently only be found again with the same hand' or 'if two different objects were placed in a participant's right and left hands then subsequently the participant could only find the object with the original hand'.

(c) This is an incorrect answer. Objects placed in the left hand could not be named. 0 marks were awarded.

**How the answer could be improved**

- If two objects are placed simultaneously, one in each hand, then removed and hidden in a scrambled pile of items / grab bag, each hand will hunt for and select its own object.
- Objects placed in one hand can only be identified through touch alone by the same hand.
- Objects put in the left hand cannot be identified in speech or writing but can be selected by the same hand from a scrambled pile of items / grab bag.

**General performance on the question**

This was a generally well answered question. There were however some candidates who did not know what 'tactile' meant, despite the term being used in Sperry's study.

## SECTION A, QUESTION 8

### Question 8

From Maguire et al's study of taxi drivers:

Describe the sample used as the experimental group in this study.

[4 marks]

## SAMPLE ANSWERS

(a) The experimental group in this study was 16 London taxi drivers. The taxi drivers had a mean age of 44 and had an age-range of 32-62 years. They were all male and also all right-handed. All of the taxi drivers had been licensed taxi drivers for at least a year and a half.

**Mark – 4 marks**

(b) There were 40 participants that were all right-handed and had been taxi drivers for over 18 months.

**Mark – 3 marks**

(c) The experimental group consisted of London taxi drivers

**Mark – 1 mark**

(d) The sample was 16 taxi drivers, all right-handed and from London and 16 non-taxi drivers also all right-handed. All 32 participants were men and the taxi drivers had been taxi drivers for at least 18 months.

**Mark – 0 marks**

## COMMENTARY

(a) This is an accurate, detailed description of the experimental group which includes reference to numerous features. The full 4 marks were awarded.

(b) Only three correct features of the experimental group have been identified: right-handed + taxi drivers + taxi drivers for over 18 months. Only 3 marks could therefore be awarded.

(c) Although the information presented was correct, this was awarded 1 mark for an extremely basic answer as there is no real description or any detail.

(d) The candidate has tried to describe the sample, and has not made it clear which is the experimental group as required by the question. This response is therefore not creditworthy and so gained 0 marks.

**How the answer could be improved**

The experimental group consisted of 16 right-handed, male, licensed London taxi drivers with a mean age of 44 years and an age range of 32-62 years. All had been licensed London taxi drivers/passed the Knowledge for more than 1.5 years. Mean time as a taxi driver was 14.3 years with a range of 1.5-42 years. All of the taxi drivers had healthy general medical, neurological and psychiatric profiles

**General performance on the question**

Another well answered question. There were however many instances where candidates misread the question and described the whole sample, including both the experimental and control groups. Candidates who did not make it clear which group was which therefore gained no marks.

## SECTION A, QUESTION 9

### Question 9

From Dement and Kleitman's study on sleep and dreaming:

Outline two conclusions that can be drawn from the findings of this study.

[4]

### SAMPLE ANSWERS

(a) Dement and Kleitman concluded that dreaming occurs in REM. This can be drawn from the finding that participants had 80% dream recall rate when woken in REM sleep. It can also be concluded that there is an association between the pattern of eye movements and the visual imagery of the dream. This can be drawn from the findings that showed when a participant's eyes had large amounts of horizontal movement, the dream depicted throwing tomatoes backwards and forwards.

**Mark** – 4 marks

(b) Eye movement correlates with what an individual is dreaming about. Individuals can recall their dreams better from REM sleep.

**Mark** – 3 marks

(c) One conclusion that can be drawn is that dreaming happens in REM. Dreams only last for a few minutes 2-7, even though they feel longer.

**Mark** – 2 marks

### COMMENTARY

(a) Two appropriate conclusions have been drawn from Dement and Kleitman's study. Although not required by the question, both have also been supported by appropriate evidence from the study. The full 2+2=4 marks were awarded.

(b) The first suggestion is an appropriate conclusion and was awarded 2 marks. The second suggestion is a partial conclusion as the comparison has not been completed so only 1 mark was awarded here. The answer was therefore awarded 2+1=3 marks.

(c) The first suggestion is an appropriate conclusion so was awarded 2 marks. The second suggestion is not a conclusion and is, in any case not something found in this study so no marks were awarded. The response therefore gained 2+0=2 marks.

**How the answer could be improved**

One conclusion that can be drawn from Dement and Kleitman's study is that dreams are more likely to be recalled if an individual is awoken whilst in REM sleep than NREM. Another conclusion could be that there is a significant association between eye movement and dream content.

**General performance on the question**

Many candidates scored well on this question. However some candidates only scored partial marks as they either gave findings rather than conclusions or failed to give two conclusions.

## SECTION A, QUESTION 10(a)

### Question 10(a)

From Piliavin, Rodin and Piliavin's subway Samaritan study:

Outline how data was recorded in this study.

[2]

## SAMPLE ANSWERS

(a) Data such as sex of helper, movement out of area, race of helper, time of help were noted down by two female observers located in the adjacent area of the subway carriage, next to the critical area where the victim collapsed.

**Mark – 2 marks**

(b) Two female observers recorded the behaviour. One was in the critical area and the other was in the adjacent area.

**Mark – 1 mark**

(c) Data was recorded quantitatively by counting how many people got on the train, how many seconds it took for someone to help and how many people went to help.

**Mark – 0 marks**

## COMMENTARY

(a) This is a clear, fully contextualised outline of how data was recorded in Piliavin's study. The full 2 marks were awarded.

(b) The second part of this response is inaccurate as both observers sat in the adjacent area. This therefore is an only partially correct response and was awarded 1 mark.

(c) This is an incorrect response as it refers to WHAT type of data is recorded, not HOW data was recorded. 0 marks were awarded here.

### How the answer could be improved

Data was gathered by two female confederates who took seats outside the critical area and recorded / noted data, through observations, for the duration of the ride, about the behaviour and characteristics of passengers such as the race of the helper and the speed of helping the victim.

### General performance on the question

This question was well answered with many clear, contextualised responses.

## SECTION A, QUESTION 10(b)

### Question 10(b)

From Piliavin, Rodin and Piliavin's subway Samaritan study:

Suggest one weakness of the way data was recorded in this study.

[2]

## SAMPLE ANSWERS

(a) One weakness is that findings could lack validity. The train could have become really crowded so the observers couldn't see clearly what was happening. They might have missed some relevant behaviours or recorded the number of people who helped the victim incorrectly. This means the data recorded is not truly reflective of people's helping behaviour.

**Mark – 2 marks**

(b) The observer may have missed important behaviours and so the study lacks validity.

**Mark – 1 mark**

(c) One weakness of conducting a covert observation such as this is that participants are unaware they are being observed and consequently the study breaks the ethical guideline of 'consent'.

**Mark – 0 marks**

## COMMENTARY

(a) The full 2 marks were awarded for a clear, fully contextualised suggestion of a possible weakness of how data was recorded in this study.

(b) Although an appropriate weakness has been identified, the response has not been contextualised to the named study so only 1 mark was awarded.

(c) 0 marks were awarded as this is an incorrect answer. It refers to a weakness of the observation method, not a weakness of how data was recorded in this study.

**How the answer could be improved**

- If the carriage where the incident occurred was particularly crowded or a passenger got in the way, the view of the observers may have been blocked so they could not see properly what was happening leading to relevant data being missed.
- Observers had so much data to record e.g. race, sex, and location of every rider seated or standing in the critical area, they may easily have missed something, so data was invalid.
- Observers could have shown observer bias e.g. by noting down that a passenger did not help – to support the diffusion of responsibility hypothesis - when in fact they may have tried to but could not get to where the victim lay, making findings invalid.

**General performance on the question**

Although this question was generally well answered, marks were often lost through lack of contextualisation to the study.

# SECTION A, QUESTION 11

## Question 11

From Reicher and Haslam's BBC prison study:

Suggest two weaknesses of recording this psychological study for television.

[4]

## SAMPLE ANSWERS

(a) One weakness of recording the study for television is that knowledge of this may lead to demand characteristics. The prisoners may have tried to aggravate the guards on purpose to make 'good television viewing'; so their behaviour would not have been natural. Another weakness is that as the participants knew this study was being recorded for television, they may have acted in a socially desirable way. For example the guards may not have acted in a harsh manner because they wanted to be seen in a 'good light' by television viewers.

**Mark** – 4 marks

(b) By having participants aware that their actions will be broadcast publicly on TV, they may experience social desirability effect and change their natural behaviour to appeal to the public. For example guards were afraid to be too harsh on the prisoners so not to appear as mean and nasty to the audience.

**Mark** – 2 marks

(c) A weakness of recording this is that the participants may have shown different unnatural behaviours, therefore affecting the validity of the study. Another weakness is that the participants' privacy is not respected therefore causing ethical issues.

**Mark** – 1 mark

## COMMENTARY

(a) Two detailed and fully contextualised suggestions have been made so the answer was awarded 2+2=4 marks.

(b) The question requires the inclusion of two weaknesses, therefore as only one appropriate and contextualised suggestion has been made only 2 marks could be awarded.

(c) One appropriate suggestion has been made but this has not been contextualised to Reicher and Haslam's study. The second suggestion is not creditworthy as participants gave informed consent before the study took place and were fully aware they would be filmed throughout the study. The response was therefore awarded 1+0=1 mark

**How the answer could be improved**

- Both guards and prisoners were aware that they would be constantly observed by television cameras and that anything they did might be shown on national television. Therefore, they might have responded with demand characteristics and acted how they thought the researchers/television viewers would want them to act. For example the prisoners may have been particularly awkward to make good television viewing.
- Both guards and prisoners were aware that they would be constantly observed by television cameras and that anything they did might be shown on national television. Therefore, they might have responded in a socially desirable way by playing their roles as they thought socially who want them to. For example the guards may have decided to be kind and thoughtful towards the prisoners on purpose so they were seen in a good light / so they were not seen to be cruel and nasty by the viewers.

**General performance on the question**

Few candidates actually contextualised their answers to the study and merely referred to 'participants'. There were also many instances that suggested candidates did not know the difference between 'social desirability' and 'demand characteristics'. Some candidates referred to ethics which was not relevant for the requirements of the question. Despite this, many candidates managed to score at least half marks on this question.

# SECTION A, QUESTION 12

## Question 12

From Milgram's study of obedience:

Describe two ways in which participants in this study were deceived.

[4]

## SAMPLE ANSWERS

(a) One way in which the participants were deceived is that they believed they were giving real electric shocks through the generator whereas the generator was completely fake and no shocks were actually given. Another way in which the participants were deceived was that they thought the study was an investigation into memory and learning rather than the true topic of obedience.

**Mark** – 4 marks

(b) One way in which the participants were deceived is that they were told that they and the other 'participant' both had an equal chance of being picked as 'teacher' or 'learner'. Another way the participants were deceived is that they genuinely thought they were administering electric shocks to the learner

**Mark** – 2 marks

(c) The participants in this study were deceived because they did not know the real aim of the study that they were taking part in. Another way the participants were deceived was that they were forced to shock the 'learner'. This made them very agitated and put them in a situation they did not want to be in. The participants were also not debriefed at the end of the study.

**Mark** – 1 mark

## COMMENTARY

(a) Two clear, fully contextualised descriptions have been provided. The candidate has identified how the participants were deceived and has explained the deception in the context of Milgram's study. The response was therefore awarded 2+2=4 marks.

(b) Two ways in which the participant was deceived have merely been identified. There is no description or explanation of the deception. Only partial marks were therefore awarded for each suggestion, making the response worth 1+1=2 marks.

(c) One way in which participants were deceived (they did not know the real aim of the study) has been identified but not described or explained in the context of the study. The second suggestion is too vague and does not relate to deception and is therefore not creditworthy. The last suggestion is incorrect as participants were debriefed after the study. The response was therefore awarded 1+0=1 mark.

**How the answer could be improved**

- One way in which participants were deceived is over the purpose of the study. They believed it was concerned about memory and learning when it was really about obedience to authority. A second way participants were deceived is that they believed they had equal chance of being teacher or learner whereas the lots were rigged to they were always the teacher and the confederate was always the learner.
- One way in which participants were deceived is that because the shock machine looked real they believed it actually gave the learner electric shocks when in reality it didn't. A second way participants were deceived is that the bangs on the wall and the shouts made by the learner in the other room led them to believe they were really hurting him, whereas in reality he was not being harmed at all.

**General performance on the question**

This was generally a well answered question, with good candidates going beyond merely giving an example of how participants were deceived, but by also showing how they were deceived.

## SECTION A, QUESTION 13

### Question 13

From Thigpen and Cleckley's study into multiple personality disorder:

Based on test results and character features, describe one of Eve's personalities.

[4]

### SAMPLE ANSWERS

(a) The character of Eve White showed her to be a caring, loving mother and wife. The inkblot test showed that Eve White had a repressive personality. The IQ test showed that she had an intelligence score of 110 and the memory test showed her to have a memory above her intelligence. Eve White was also described to be quite shy and not particularly bold or confident.

**Mark** – 4 marks

(b) Eve Black was mischievous and flirty. She was outgoing and loved to go out and have fun. She was regressive but healthier than Eve White. She had a lower IQ and poorer memory than Eve White. She was completely different to Eve White and took absolutely no responsibility for her child or husband.

**Mark** – 3 marks

(c) One of Eve's personalities was Eve Black who had an extrovert personality. She was very aggressive and denied being married with a child even though Eve White was. She also despised her husband and expressed violent feelings towards him. Eve Black's IQ score was 104 whereas Eve White's was 110 and Eve Black on the memory test was shown to have a poor memory.

**Mark** – 2 marks

(d) Eve White was one personality. She had an IQ of 110.

**Mark** – 1 mark

### COMMENTARY

(a) This is an accurate description of Eve White which refers to both character features and psychological tests. The full 4 marks were awarded.

(b) Although the candidate has referred to both character features and psychological test results, both are rather vague and could have been developed further by providing some fine details. The response was however felt to be worth 3 marks.

(c) Much of this response does not relate to character features or test results. The reference to Eve Black having a poor memory lacks clarity – her memory was on par with her IQ though it was poorer than Eve White's which was above her IQ. The response was awarded 2 marks.

(d) This is an extremely basic answer with reference merely made to one psychological test and no references at all to character features. 1 mark was awarded.

### How the answer could be improved

- Eve White: was found to be demure, retiring, neat, colourless, conservative, honest, serious, conscientious, but anxious. She was seldom animated. Psychological tests showed her to have an IQ of 110, a memory far above her IQ, and to be repressive. A physiological test of an EEG showed her to have alpha rhythms of 10½-11½ cycles per second.
- Eve Black: was found to be a party girl, mischievous, childish, egocentric, vain, amusing, likable and provocative. Psychological tests showed her to have an IQ of 104, a memory on the same level as IQ. She was also reported to be repressive. A physiological test of an EEG showed her to have alpha rhythms of 12-13 cycles per second.
- Jane: was found to be mature, vivid, sensible, capable, interesting and impressive. The physiological test of an EEG showed her to have alpha rhythms of 10½-11½ cycles per second, the same as Eve White.

### General performance on the question

Although some candidates provided somewhat muddled answers which referred to both Eve White and Eve Black, there were many excellent responses which gave accurate details in relation to both test results and character features of one named personality.

## SECTION A, QUESTION 14

### Question 14

From Griffiths' study into fruit machine gambling:

Describe how qualitative data was gathered in this study.

[4]

### SAMPLE ANSWERS

(a) In Griffiths' study qualitative data was gathered from interviews after the participants had played on the fruit machines. Some information gathered included thoughts about the participants' own level of skill at gambling and whether or not they considered themselves of average, above or below average skill. Other qualitative data was gathered by Griffiths whilst the participants spoke aloud in the speaking aloud condition where comments made by the RGs and NRGs were recorded via a lapel microphone onto a tape recorder.

**Mark – 4 marks**

(b) The speaking aloud group had to say everything they thought. They were told to keep talking and not to censor their thoughts. There was a speaking aloud group for both the regular and non-regular gamblers. The verbalisations were recorded using a tiny microphone and were transcribed within 24 hours of completing the experiment.

**Mark – 3 marks**

(c) Qualitative data was gathered through having participants think aloud in the 'thinking aloud condition'. As they played on the machines they spoke into a microphone all that they were thinking and feeling e.g. they made comments like, "This machine likes me" when they won and said swear words when they lost.

**Mark – 2 marks**

(d) Qualitative data was gathered through asking them questions (an interview) about gambling.

**Mark – 1 mark**

### COMMENTARY

(a) The candidate has referred to both ways in which qualitative data was gathered in this study. The response is accurate and adequately detailed to warrant awarding the full 4 marks.

(b) This is a fairly accurate description of the way qualitative data was gathered. The candidate has however failed to state that the microphone was (a) a lapel microphone or (b) that it was connected to a (portable) tape recorder. There is also no reference to the post-experimental semi-structured interviews. Despite this, because the candidate has referred to the data being transcribed within 24 hours, there is adequate information to gain 3 marks.

(c) As the candidate has only referred to the thinking aloud condition and the use of microphones, this is a rather vague answer that lacks real detail. It is however adequate for 2 marks.

(d) As this is a mere identification of one way in which qualitative data was gathered so the response was only awarded 1 mark.

### **How the answer could be improved**

Qualitative data was gathered through self reports made firstly by RGs and NRGs in the 'thinking aloud' condition. Participants in this condition had all their verbalisations recorded using a lapel microphone connected to a portable tape recorder. All verbalisations made during the gambling session of each participant were transcribed within 24 hours. In addition, qualitative data was gathered in the post-experimental semi-structured interview in which participants were asked a number of skill-related questions.

### **General performance on the question**

Although many candidates were able to refer to the 'thinking aloud' condition, many thought the qualitative data was gathered through observation and failed to mention the use of a (lapel) microphone and tape recorder. Only the best candidates were able to refer to both how the data was gathered through the 'thinking aloud' condition and the post-experimental interviews..

## SECTION A, QUESTION 15

### Question 15

From Rosenhan's study 'On being sane in insane places':

Describe two pieces of evidence supporting the suggestion that pseudopatients were depersonalised.

[4]

### SAMPLE ANSWERS

(a) From Rosenhan's study, the pseudopatients reported that there was lack of privacy which made them feel depersonalised. For example, in some hospitals, the toilets did not have any doors on them. Another piece of evidence comes from how the staff and psychiatrists responded and interacted with the pseudopatients making them feel depersonalised. For example, many of the staff did not make eye contact or acknowledged the pseudopatients when they were asked

**Mark** – 4 marks

(b) When the pseudopatients attempted to talk to the staff they were frequently ignored, making them feel as though they didn't exist. Also the pseudopatients had no privacy when using the toilets as they had to keep the door open.

**Mark** – 3 marks

(c) The pseudopatients were often ignored when they asked the hospital staff questions. All of their behaviour was seen as abnormal due to the stickiness of the label of schizophrenia or bipolar. For example, taking notes was seen as writing behaviour.

**Mark** – 2 marks

(d) The toilets had no doors and staff made no eye contact with the pseudopatients.

**Mark** – 2 marks

### COMMENTARY

(a) Two accurate and fully contextualised descriptions have been provided so the full 2+2=4 marks were awarded.

(b) The first suggestion is accurate and fully contextualised but the second suggestion is not totally accurate – pseudopatients did not report having to keep the toilet doors open, they reported that some toilets had no doors. The response was therefore awarded 2+1=3 marks.

(c) The first suggestion is accurate and fully contextualised but the second suggestion, as given here, does not show how pseudopatients felt depersonalised, it refers more to how their behaviour was misinterpreted. The response was therefore awarded 2+0=2 marks.

(d) Two ways in which the pseudopatients were depersonalised have merely been identified. Neither piece of evidence has been fully described or elaborated on. The response was therefore awarded 1+1=2 marks.

### **How the answer could be improved**

- One way the pseudopatients were depersonalised is that the toilets in the hospitals often had no doors on them so they had no privacy. Another way they were depersonalised is that on admittance several pseudopatients underwent initial physical examinations in semi-public rooms. In these rooms, staff members went about their own business as if the pseudopatients were not there, or did not exist.
- One way the pseudopatients were depersonalised is that on one occasion a nurse did not seem to acknowledge the patients' presence and unbuttoned her uniform to adjust her bra in the presence of both a pseudopatient and other genuine patients. Another way they were depersonalised is when pseudopatients tried to engage in conversations with staff or ask staff questions, they were usually ignored making them feel as if they didn't exist

### **General performance on the question**

Overall, another well answered question. However, there were a number of instances where candidates lost marks because they did not make it clear how their evidence specifically related to depersonalisation.

## SECTION B, IN RELATION TO DEMENT & KLEITMAN – SLEEP AND DREAMING, QUESTION 16(a)

### Question 16(a)

Outline why your chosen study was conducted.

[2]

## SAMPLE ANSWERS

(a) Dement and Kleitman's study was conducted to test whether there are links between REM and dreaming, whether the direction of eye movement correlates with dream content and whether dreams occur during REM or NREM sleep.

**Mark – 2 marks**

(b) To find out if more dreams occurred during REM sleep than NREM sleep, if there was a correlation between eye movement and dream content, and if participants were correctly able to say whether they had been dreaming for 5 or 15 minutes.

**Mark – 2 marks**

(c) The study was conducted to see whether dream recall and dream duration estimations changed when in REM and NREM sleep.

**Mark – 1 mark**

(d) Dement and Kleitman's sleep and dreaming study was conducted to see when we dream and to further investigate REM and NREM cycles.

**Mark – 1 mark**

## COMMENTARY

(a) A clear and contextualised reason for why the study was conducted. The response was awarded the full 2 marks.

(b) The full 2 marks were awarded here for a clear and contextualised reason for why the study was conducted.

(c) This is vague/partial response which is not actually very accurate i.e. participants did not have to estimate how long they thought they had been dreaming in NREM sleep. Some material is worth crediting, so the response was awarded 1 mark.

(d) This is a vague response and does not actually identify the main aim of the study: to investigate the relationship between REM sleep and dreaming. It was therefore awarded 1 mark.

**How the answer could be improved**

- The aim of Dement and Kleitman's study was to investigate the relationship between REM sleep and dreaming.
- The aim of Dement and Kleitman's study was to observe and record the length, frequency and patterns of the participants' REM sleep to discover if there is any relationship between these and dreaming.

**General performance on the question**

This was well answered. Good responses were clear and fully contextualised. Weak responses omitted important material, e.g. failing to state that the study was investigating the relationship between REM and dreaming.

# SECTION B, IN RELATION TO DEMENT & KLEITMAN – SLEEP AND DREAMING, QUESTION 16(b)

## Question 16(b)

Describe how the self-report method was used to gather data in your chosen study.

[4]

## SAMPLE ANSWERS

(a) Participants were woken by a bell next to their bed in either REM or NREM. They spoke through a tape recorder. Participants had to say whether they had been dreaming or not. If they were they had to state how long their dream was (estimate) – 5 or 15 minutes. They then had to describe their dream content. Sometimes the researchers came into the sleep lab to ask further questions, otherwise they could go back to sleep.

**Mark – 4 marks**

(b) When participants were woken by a doorbell from either REM or NREM sleep, they were told to talk into a voice recorder next to their beds. They were first told to state whether or not they had been dreaming and then, if possible, to recall the content of their dream.

**Mark – 3 marks**

(c) Self-report was a method used in this study. Participants were woken up at various times while sleeping and asked to talk through and describe any dream they had through a recording device.

**Mark – 2 marks**

When the participants were woken up they immediately had to talk into a recorder and tell if they had been dreaming and the content of their dream. They did this without talking to the researcher, so they were not told anything when they woke up that could affect what they said into the recorder. They just did it as soon as they were woken.

**Mark – 2 marks**

## COMMENTARY

(a) This is a clear and accurate description of how the self-report method was used in Dement and Kleitman's study, with reference to both the use of the tape recorder and the researcher to gather data. The full 4 marks were awarded.

(b) Not quite enough depth is included to be awarded full marks. There is no reference to that some participants had to estimate how long they had been dreaming out of 5 or 15 minutes or that, on occasions, the experimenter entered the room to question them further on a particular point of their dream. 3 marks were therefore awarded.

(c) This was awarded 2 marks as it is not a fully developed response. There is no reference to how the participants were awoken, that it was sometimes during REM and sometimes during NREM, what participants were specifically asked to report about, or that, on occasions, the experimenter entered the room to question them further on some particular point of their dream. The candidate has however referred to the use of a tape recorder and that participants had to describe their dream once they were woken up.

(d) This was awarded 2 marks as it is not a fully developed response. There is no reference to how the participants were awoken, that some had to estimate how long they had been dreaming out of 5 or 15 minutes or that, on occasions, the experimenter entered the room to question them further on some particular point of their dream.

### **How the answer could be improved**

Participants were woken whilst in either REM or NREM sleep. For all awakenings, they spoke into a recording device near their bed. They were instructed to first state whether or not they had been dreaming and then, if they could, to relate the content of their dream. When the participant had finished speaking, the experimenter, who could hear their voices, occasionally entered the room to further question them on some particular point of the dream.

### **General performance on the question**

Here only the more able candidates appreciated that the self-report method was used to gather data in at least two ways in their chosen study e.g. talking into a tape recorder and the experimenter entering the room to ask additional questions/clarify points. Other strong candidates were able to describe in depth one way the self-report method was used.

# SECTION B, IN RELATION TO DEMENT & KLEITMAN – SLEEP AND DREAMING, QUESTION 16(c)

## Question 16(c)

Describe one strength and one weakness of gathering data using self-reports. Support your answer with evidence from your chosen study. [6]

## SAMPLE ANSWERS

(a) A strength of using self-reports especially in relation to this study is that they give the individuals' perception of their internal mental thinking processes. This allowed the participants to tell the researchers the contents of their dreams as otherwise the researchers have no other way of finding out these internal thinking processes.

However a weakness of self-reports is that the participants could lie in order to give the answers they believe the researcher wants to record and this affects the reliability and validity of the findings. For example the participants may lie about whether they had been dreaming or not and about the content of their dreams.

**Mark – 6 marks**

(b) One strength is that the participants can write/say/describe their experience in their own words. You then get in depth qualitative data. In Dement and Kleitman's study they immediately talked about their dreams into the recorder, without being told anything from the researcher, so you know it's subjective.

One weakness is that they know they are taking part and that people will look at their results. They therefore may want to show themselves in a good light or give results that they believe are right.

**Mark – 3 marks**

(c) A strength of gathering data using the self-report method is that it is quick to do as there is no IV to manipulate, participants are just asked to state whether or not they dreamed and then recall their dream.

A weakness is that the participant may lie about what they dreamt about – for example, if they had dreamed about something personal or embarrassing, they may not want to report it and instead make up a dream. This would affect the results, decreasing their validity.

**Mark – 3 marks**

(d) A strength of using self-report is that it gives the data on a personal first-hand level so there is no researcher bias that could be given if it wasn't self-report – it allows the participant to give an accurate and detailed description.

A weakness however would be that there could be social bias and some of the answers given may not represent how they actually feel but seem like the socially accepted way to think. Therefore not giving an accurate representation.

**Mark – 2 marks**

## COMMENTARY

(a) Strength: identified (they allow an individual's thoughts to be accessed), elaborated through the reference to the strength's implication (no other way for researchers to access an individual's thoughts), contextualised through the reference to dream content. Weakness: identified (participants can lie), elaborated through the reference to the weakness' implication (findings lack validity), contextualised through the suggestion that participants may have lied about whether or not they had been dream and their dream content. Both the strength and weakness have been appropriately identified, explained and contextualised so the response was awarded 3+3=6 marks.

(b) Strength: 1 mark for an appropriate strength (participants can describe their experiences in their own words), 1 mark for the justification/elaboration (in-depth qualitative data can be gathered), no actual supporting evidence = 2 marks. Weakness: 1 mark for the identification of an appropriate weakness (social desirability/demand characteristics), no elaboration and no evidence. The response was therefore awarded 2+1=3 marks.

(c) Strength: No marks awarded as data can be gathered quickly in other research methods e.g. observation. Also the reference to not having to manipulate IVs is irrelevant to the question. Weakness: 1 mark for identifying an appropriate weakness (participants may lie), 1 mark for elaboration through the reference to the weakness's implications (decreasing validity) + 1 mark for the contextualisation (if they had dreamed about something personal or embarrassing, they may not want to report it and instead make up a dream). The response was therefore awarded 0+3=3 marks.

(d) Although an appropriate strength (personal, first-hand, accurate and detailed description) and an appropriate weakness (answers may not reflect what the participant actually feels as they have responded in a socially desirable way), neither suggestion has been contextualised in any way. The response can therefore only be awarded 1+1=2 marks.

### How the answer could be improved

A strength of self-reports is that they can generate large amounts of quantitative and qualitative data to give a comprehensive overview of the participant. Here Dement and Kleitman were able to gather quantitative data by recording how many times participants could recall dreams when awoken from either REM or NREM and qualitative data in relation to dream content. For example, one participant dreamed of throwing tomatoes.

A weakness of self-reports is that responses may have a social desirability bias because participants give an answer which they think makes them look good or is acceptable to society. For example, if a participant was having a particularly embarrassing dream, they may not wish to report this and so say they were dreaming about something different. This will affect the validity of the findings as they may not be true.

### General performance on the question

Many candidates were able to describe, explain/justify and support with evidence both an appropriate strength and an appropriate weakness of the self-report method. Some candidates however failed to support their answers with adequate evidence, while others gave study-specific answers, which was not required.

# SECTION B, IN RELATION TO DEMENT & KLEITMAN – SLEEP AND DREAMING, QUESTION 16(d)

## Question 16(d)

Outline the findings of your chosen study.

[8]

## SAMPLE ANSWERS

(a) It was found that when in REM sleep people are more likely to remember a dream they had/were having than if they were in NREM sleep. They also found that people were able to correctly estimate how long they had been dreaming for (e.g. participants were asked 'Have you been dreaming for 5 or 15 minutes?') The results were that participants dreaming for 5 minutes were correct 45/51 times and participants dreaming for 15 minutes were correct 47/60 times. It was also found that the direction of eye movement (none/horizontal/vertical) corresponded to the content of the dream (e.g. no eye movement = dreaming about driving, horizontal movement = dreaming about throwing tomatoes, vertical movement = dreaming about climb a ladder and looking up and down.

**Mark** – Top band, 7-8 marks

(b) In this study of sleep and dreaming, one finding of the study was that during the study 152 dreams were recalled during REM sleep, whereas only 11 were recalled during NREM sleep, and only the latter part of the dream could be recalled. This shows association between REM sleep and dreaming. Another finding was that eye movements were shown to be related to dream content e.g. a participant dreamt of watching two people throwing tomatoes at each other, and this participant was also recorded to have made horizontal eye movements during sleep. Another finding made was that all participants went through REM during the stages of sleep and only one participant recalled a dream during NREM and only the latter part (thought to be a memory from REM sleep). Another finding was that during REM sleep fast EEG activity was recorded for each participant.

**Mark** – Mid band, 4-6 marks

(c) 1. It showed that more dreaming tended to occur during REM sleep than non-REM sleep. This is supporting the hypothesis that there is a correlation between them.  
2. Participants, when woken up 5 or 15 minutes after REM had occurred were in most cases able to correctly match/say their time. This is also supporting the hypothesis showing that we dream in real time.  
3. Participants were woken up after 1 of 4 eye movements had occurred for a minute and the results were supporting the hypothesis that there is a correlation between eye movement and dream content in this study.  
All of Dement and Kleitman's hypotheses were supported.

**Mark** – 3 marks

(d)

- They found that 80% of the participants remembered what happened in their dream when in the state of REM sleep.
- Only a few participants were able to identify how long they had been in REM sleep for.
- Some participants were not able to explain the content of the dream even when in REM which showing that some participants may be showing social demands as their dream may be inappropriate.

**Mark** – Bottom band, 1-3 marks

## COMMENTARY

(a) Description of results is mainly accurate and several fine details have been included. Detail is appropriate to the level and time allowed and the answer shows good understanding. The response was placed in the top band.

(b) The description of the results is mainly accurate though the reference to 'only one participant recalled a dream during NREM and only the latter part (thought to be a memory from REM sleep)' is incorrect. There are several key omissions, but understanding is clearly evident. The response was therefore placed in the middle band.

(c) The description of results is very basic and no fine details have been included. The answer is very list-like and shows little real understanding. It was placed in the bottom band and awarded 3 marks.

(d) There are a number of inaccuracies in this response - Result 1: incorrect as there was 80% recall in REM, not 80% of participants could recall dreams from REM; result 2: incorrect as all participants could recall dreams from NREM though fewer were recalled from NREM than REM; result 3: incorrect as participants were, overall, good at being able to identify how long they had been in REM (92/101 correct estimates over the two times); result 4: poorly expressed but as on some occasions participants couldn't recall dreams from REM (e.g. DN couldn't 9/17 times). There is some creditworthy material here, but most of the findings are incorrect. The answer was therefore placed in the bottom band.

### How the answer could be improved

Overall, more dreams were recalled from REM (152/191) than NREM (11/160). When participants were asked to estimate how long they had been dreaming, there were more correct estimates (92) than incorrect (19). However there was a lower percentage of correct estimates of dream length when participants were awoken after 15 minutes of REM (78%) than 5 minutes of REM (88%). The length of dream narrative was found to have a significant relationship to the duration of the REM period before awakening. Findings also showed there to be a strong association between the pattern of eye movements in REM and dream content. For example, a participant whose eyes were moving horizontally reported dreaming of watching two people throwing tomatoes at each other, a participants whose eyes were moving vertically reported dreaming about throwing basketballs at the net, first shooting and looking at the net, and then looking down to pick another ball off the floor. Participants who were woken after a period of no or little eye movement typically reported watching something in the distance or staring fixedly at an object.

### General performance on the question

This question saw the full range of marks being awarded. Some answers were very muddled and 'waffly', although there was evidence that many candidates knew this study really well.

# SECTION B, IN RELATION TO DEMENT & KLEITMAN – SLEEP AND DREAMING, QUESTION 16(e)

## Question 16(e)

Suggest improvements to your chosen study.

[8]

## SAMPLE ANSWERS

(a) One problem with Dement and Kleitman's study is that it is low in ecological validity. It is not normal for participants to be sleeping in a sleep lab. To improve this, I would conduct the experiment in the participant's own home, in their own bedroom and bed. The researcher would come and spend the night at the participant's house each night of the experiment. We would still use the EEG and EOG method like in the original study to measure brain wave activity and REM patterns. They would also report for bed at their normal bedtime and would be allowed to drink caffeine and alcohol if they normally did.

Another problem with the study is that it has a small and limited sample, therefore findings cannot be generalised to other people's REM patterns and dreaming. Also there is gender imbalance. To improve this I would increase the sample size from 9 participants (2 females, 7 males) to 30 participants, all from different countries. For example England, China, India, Zambia, Russia. I would also have one male and one female from each country to make sure there is a gender balance. I would choose participants from various occupations and backgrounds

**Mark** – Top band, 7-8 marks

(b) One way to improve Dement and Kleitman's study would be to increase the sample size from 9 participants to 50. Also use an equal amount of males and females e.g. 25 males, 25 females, instead of 7 males and 2 females. Furthermore the researcher should involve other countries and cultures to increase the study's representativeness and generalisability.

Another improvement could be to the study's ethics. Dement and Kleitman should decrease their deception by informing the participant whether or not they were in REM sleep. They could also make it easier for the participant to withdraw as they may have found it stressful whilst being observed.

A further improvement to the study would be to the procedure. The controls could have been altered so that the participant would be able to drink their usual amount of caffeine or alcohol. Also the experiment could take place in the participant's home instead of a lab.

**Mark** – Third band, 5-6 marks

(c)

- It was a small sample group. It was only 9 participants. 5 of them were studied intensely and the other 4 just to confirm the results. They should have used a bigger sample group.
- Only 2 of the 9 participants were females. It should be an equal number.
- The study is low in ecological validity. To sleep in a lab and with electrodes (EEGs) may have affected their normal sleep patterns. It is not a normal environment. If they did the study in the participant's own house that would be better.
- They didn't have any control over what the participants did earlier in the day. They were only asked to avoid alcohol and caffeine because they are known to affect people's sleeping patterns but did they have enough control over other possible variables before the participants got to the lab? They should have been asked if there were possibly other things during the day that could affect their normal sleeping patterns, like stress.

**Mark** – Second band, 3-4 marks

(d) I would improve the study by conducting it in the participant's own home. I would increase the sample size and have an equal number of males and females and make sure participants came from more than one country. I would improve the ethics of the study by making sure all participants were debriefed afterwards.

**Mark** – Bottom band, 1-2 marks

## COMMENTARY

(a) Two appropriate improvements have been suggested and described in detail. There is good understanding and expression is clear. Improvements have been justified and implementation has been considered, i.e. the candidate has said 'why' and 'how' the suggestions would be improvements. As a result the response was placed in the top band.

(b) Several appropriate improvements have been suggested, although no obvious thought has been given as to how to implement the first suggestion, or really the last suggestion. Some suggestions haven't been justified e.g. 'the experiment could take place in the participant's home instead of a lab'. There are some appropriate links to the chosen study, so the response was placed in the third band.

(c) The first three suggestions are very basic with little or no real consideration as to how they might be implemented. The suggestion relating to a lack of control over what participants did during the day doesn't actually lead anywhere. The answer was therefore placed in the second band.

(d) The description is extremely basic and lacks any detail. No thought has been given as to how any of the suggestions could be implemented, or why the suggestions are improvements. There is no contextualisation, so the response could apply to any of the named studies. The answer was therefore placed in the bottom band.

### How the answer could be improved

One problem with Dement and Kleitman's study is that it is low in ecological validity. It is not normal for participants to be sleeping in a sleep lab. To improve this, I would conduct the experiment in the participant's own home, in their own bedroom and bed. The researcher would come and spend the night at the participant's house each night of the experiment. We would still use the EEG and EOG method like in the original study to measure brain wave activity and REM patterns. They would also report for bed at their normal bedtime and would be allowed to drink caffeine and alcohol if they normally did.

I could also improve Dement and Kleitman's study by having a more representative sample. This study only had 9 participants, 7 men and 2 women. This is not only a small sample but also contains a male gender bias. I would therefore increase the sample size to 40 (20 males and 20 females). I would gather my sample by putting an advertisement in a local newspaper asking for volunteers between the ages of 16 and 60 to take part in a study of sleep and dreaming. Once all applications have been received I would select my sample, making sure I had an equal number of males and females and that I had a good age and occupation range.

### General performance on the question

A good range of marks was awarded for this question. Many candidates were able to suggest and justify two or more appropriate improvements. Real understanding was shown by those candidates who both explained why their suggestions would improve their chosen study and described how they would implement their suggestions. There were some instances where answers lacked contextualisation e.g. I would improve the sample by having more participants; I would improve the ecological validity by conducting the study in a natural environment.

## SECTION B, IN RELATION TO DEMENT & KLEITMAN – SLEEP AND DREAMING, QUESTION 16(f)

### Question 16(f)

16(f). Evaluate the improvements you have suggested to your chosen study.

[8]

### SAMPLE ANSWERS

(a) A good thing about making the study more ecologically valid is that the researchers will obtain more realistic results. For example, the participants will be in their own bedroom, and will drink caffeine and alcohol normally if they did, therefore their dreaming patterns won't be affected and will be normal. This is a strength as it increases the validity and usefulness of the results to show a link between REM patterns and dreaming. A weakness is that the setting will not be controlled as in the original study (sleep lab). There will therefore be factors such as noise from road traffic, children etc that may affect sleeping and dreaming. This means the reliability of the results will be reduced.

A strength of increasing the sample is that the results will be more generalisable to other people's REM patterns and dreaming content. For example, there will be a wide range of participants from different occupations and countries. This increases the usefulness of the results because they may show that a link between REM sleep patterns and dreaming is universal. A weakness of this is that all participants may have different levels of stress and different experiences that affect dreaming. This reduces the generalisability of the results

**Mark** – Top band, 7-8 marks

(b) One strength of making this into a field experiment in the participant's own home is that the ecological validity will be improved. For example, participants still go to bed to sleep at the usual bedtime, but each participant is also studied in their own natural environment, in their house where they go to sleep in their own beds where they usually sleep every night in real life. This therefore allows the results of the study to be more generalisable to real life and so more useful. One weakness of this improvement is that there will therefore be less control over extraneous variables which could affect the results of the study. For example, variables such as light, noises, temperature could be different in each home for each participant and therefore also may affect their sleeping patterns or dream content and disrupt sleeping. This is a weakness because this decreases the validity as a true cause and effect may not be established. One effect on the findings I'd expect to see is that more dreams may be recalled during REM sleep as they are sleeping in their natural environment. Another effect I'd expect to see is that noise or light may affect sleeping patterns for different participants so fewer details of dreams could be recorded.

**Mark** – Third band, 5-6 marks

(c) An advantage of increasing the sample is that any results will be more representative. However carrying out the experiment on 50 people would be time-consuming and costly. Also, it may be difficult to find 50 people who are prepared to do the experiment.

An advantage of conducting the experiment at home would be that it would make the participants more comfortable and a better indication of their sleep patterns. A disadvantage is that this would in turn be difficult to control and other factors such as noise may affect sleep.

**Mark** – Second band, 3-4 marks

(d) By conducting the study in the participant's own home it would be more ecologically valid. Increasing the sample size and having the same number of males and females would make the sample more representative and the findings more generalisable. Also by drawing participants from different cultures would also make the sample more representative and the findings more generalisable.

**Mark** – Bottom band, 1-2 marks

## COMMENTARY

(a) Both improvements have been evaluated, giving both a strength and a weakness of the suggested improvements. The evaluation is in adequate detail for the time and level allowed. There is good understanding and the answer is competently structured and organised. The response was placed in the top band.

(b) The candidate has made an appropriate evaluation of one possible improvement (increasing ecological validity). Both strengths and weaknesses of the improvement have been considered and some thought has to be given to how the improvement may impact the results. Understanding is evident and the answer is well-linked to the chosen study. However, as only one improvement has been considered the response cannot gain more than 6 marks – see mark scheme. The response could only therefore be placed in the third band.

(c) Although the evaluation of the improvements is appropriate, it is basic and lacks any real detail. The first evaluation point is not contextualised and could therefore apply to any study, while the second evaluation point is only loosely linked to the chosen study. The response was however placed in the second band.

(d) The evaluation is extremely basic, lacks any detail and shows no real understanding in relation to the chosen study. There is no contextualisation, so the response could apply to any of the named studies. The answer could therefore only be placed in the bottom band.

### How the answer could be improved

A positive thing about making the study more ecologically valid is that the researchers will obtain more realistic results. For example, the participants will be in their own bedroom, and will drink caffeine and alcohol as normal, therefore their dreaming patterns are likely to be normal as they won't be affected by either an artificial environment or a changed life style. This is a strength as it increases the validity and usefulness of the results to show a link between REM patterns and dreaming. A negative thing with improving the ecological validity is that the setting will not be as controlled as in the original study which took place in a sleep laboratory. There will therefore be factors such as variations in noise from road traffic, noises from other people or animals, bedroom lighting and temperature that may affect sleeping and dreaming. The reliability of the results may therefore be reduced.

A good thing about having a larger and more representative sample is that findings will be more generalisable. This increases the usefulness of the results because they may show that a link between REM sleep patterns and dreaming is universal. A limitation of this improvement is that it may be difficult to get a more representative sample. There may be a large gender or occupational bias in those who volunteer to take part so I cannot get an equal number of males and females representing a wide range of occupations. Therefore any findings in relation to patterns of sleep and dreaming and associations between REM and NREM will still be limited and lacking in generalisability.

### General performance on the question

Many candidates, although able to evaluate their improvements failed to contextualise their answers to their chosen study e.g. by increasing my sample size and including both males and females I will get a more representative sample, so I can generalise my findings to the whole population. Such a statement could apply to any of the three studies offered in this section. Furthermore, many evaluation points were extremely basic and failed to show any real understanding e.g. having an equal number of males and females would make findings more generalisable.

## SECTION B, IN RELATION TO THIGPEN AND CLECKLEY – MULTIPLE PERSONALITY DISORDER, QUESTION 16(a)

### Question 16(a)

Outline why your chosen study was conducted.

[2]

## SAMPLE ANSWERS

(a) Thigpen and Cleckley's study was conducted to document the case of a 25-year old married woman called Eve White who was referred to psychiatrists due to blinding headaches and

**Mark – 2 marks**

(b) Thigpen and Cleckley gained an interest into multiple personality disorder and wanted to document therapy for it.

**Mark – 1 mark**

## COMMENTARY

(a) This is a clear and fully contextualised outline of why the study was conducted, so the full 2 marks were awarded.

(b) A very vague answer which has little real contextualisation or explanation of why the study was conducted. 1 mark was awarded.

### How the answer could be improved

- Thigpen and Cleckley's study was conducted so the psychotherapeutic treatment of a 25 year old woman who presented with a history of severe headaches and blackouts but was later discovered to have multiple personality disorder could be documented.
- Thigpen and Cleckley's study wanted to provide an account of an individual thought to have MPD and offer possible treatment.

### General performance on the question

This was well answered. Good responses were clear and fully contextualised. Weak responses included important omissions e.g. failing to include a reference to treatment when saying the study was conducted because Eve White presented with a history of severe headaches and blackouts.

## SECTION B, IN RELATION TO THIGPEN AND CLECKLEY – MULTIPLE PERSONALITY DISORDER, QUESTION 16(b)

### Question 16(b)

Describe how the self-report method was used to gather data in your chosen study.

[4]

### SAMPLE ANSWERS

(a) Over 100 hours of interviews gathered information from Eve White, Eve Black and Jane. This information was supported by her family to back up Eve White's accounts of events from childhood to adulthood. Researchers also used psychometric tests to measure both Eve White's and Eve Black's IQ, memory. The inkblot test was also conducted on the two personalities and hypnosis was induced to reveal different personalities and restore memory.

**Mark** – 4 marks

(b) Thigpen and Cleckley recorded Eve White and conducted interviews over 100 hours over 14 months with Eve White, Eve Black and Jane, while conducting personality tests, intelligence tests and memory tests on each of them to see if they were really different. This gave us both quantitative and qualitative data.

**Mark** – 3 marks

(c) Thigpen and Cleckley interviewed Eve White and her two other personalities, Jane and Eve Black, for approximately 100 hours. They would write notes during the interview from memory.

**Mark** – 2 marks

(d) Interviews were conducted over a 7 month period, totalling 32 hours between Eve and Thigpen and Cleckley.

**Mark** – 1 mark

### COMMENTARY

(a) A clear, accurate and detailed description of how the self-report method was used in this study. Reference has been made to several ways the self-report method was used and the answer has been well contextualised. The response was placed in the top band and awarded the full 4 marks.

(b) The candidate has referred to the interviews with all three personalities. However, the personality, intelligence and memory tests were not conducted on all three personalities, so this is not a completely accurate response, therefore 3 marks were awarded.

(c) This was awarded 2 marks as although all three personalities have been included and reference has been made to the use of interviews, there is no real depth or elaboration in the response.

(d) There are numerous inaccuracies in this response e.g. the length of the study and the number of interview hours. However, as the candidate has referred to the use of interviews, the response was awarded 1 mark.

**How the answer could be improved**

Over a period of 14 months, during a series of interviews, totalling approximately 100 hours, extensive material was obtained about the behaviour and inner lives of Eve White, Eve Black and Jane. Self report data was also gathered when Eve White was hypnotised. Both of Eve White's parents and her husband were also interviewed. Self-report data was additionally gathered through the psychological tests undertaken by both Eve White and Eve Black. These included an IQ test, a memory test, the Rorschach (inkblot) test and the drawings of human figures test.

**General performance on the question**

Here only the more able candidates appreciated that the self-report method was used to gather data in at least two ways in their chosen study e.g. interviews, hypnosis, psychometric testing, or were able to describe in depth one way the self-report method was used.

## SECTION B, IN RELATION TO THIGPEN AND CLECKLEY – MULTIPLE PERSONALITY DISORDER, QUESTION 16(c)

### Question 16(c)

Describe one strength and one weakness of gathering data using self-reports. Support your answer with evidence from your chosen study. [6]

## SAMPLE ANSWERS

(a) A strength of using a self-report is that it allows insight into one's thoughts and feelings which otherwise would not be found out. In Thigpen and Cleckley's study, Eve White revealed that she was hearing voices which then led to the revealing of Eve Black one of her other personalities. By interviewing Eve White, Thigpen and Cleckley could record that her thoughts consisting of hearing voices were due to another personality, so they collected useful data on MPD.

A weakness is that the participant who is reporting could be dishonest in their reports, so findings would be false. Eve White could have been acting and pretending to have other personalities, so data on her personalities and possible evidence for MPD from the interviews could be false and therefore results are not valid.

**Mark – 6 marks**

(b) A weakness of gathering data through self-reports is that people may display socially desirability characteristics e.g. in Thigpen and Cleckley, Eve White may have lied about her actions in order to save her from embarrassment. This therefore makes the results regarding the existence of multiple personality disorder less valid.

A strength of self-reports is that they allow you to gather large amounts of data directly from the participant e.g. in Thigpen and Cleckley they were able to gather information from all three personalities and Eve White's family to support Eve's stories. This makes the results regarding the existence of MPD more valid as there are large amounts of data to support the findings.

**Mark – 5 marks**

(c) One strength is that you get lots of rich, detailed data straight from the participant. For example, in this study they received information from Eve about her home life and symptoms during 100 hours of interviews.

A weakness is that people may lie to be seen as socially desirable or to respond to demand characteristics. For example, Eve Black was known to be mischievous and a liar so she may have lied during the interviews giving incorrect information.

**Mark – 4 marks**

(d) One strength of using self-report is that you gain both qualitative and quantitative data. With Eve White we got her thoughts and feelings through interviews to learn what was going on, this is qualitative data. Also quantitative data from tests which is easy to analyse.

One weakness is that there may be observer bias where the observer makes it out that Eve does have multiple personalities and they did begin to like Eve Black who was flirtatious and may have acted differently for Thigpen and Cleckley showing demand characteristics.

**Mark – 2 marks**

## COMMENTARY

(a) Strength: 1 mark for identifying an appropriate strength (insight into one's thoughts and feelings), 1 mark for elaboration through the reference to the strength's implications (collected useful data on MPD) + 1 mark for the contextualisation: Eve White thought she was hearing voices which was due to her other personalities = 3 marks. Weakness: 1 mark for identifying an appropriate weakness (possibility of dishonesty in responses), 1 mark for elaboration through the reference to the weakness's implications (results will not be valid) + 1 mark for the contextualisation: 'pretending to have other personalities' = 3 marks. The response was therefore awarded 3+3=6 marks.

(b) Weakness: identified (social desirability), explained/justified (makes findings in relation to MPD less valid), however the evidence from the study is very vague and could really apply to any of the studies named in this section. (2 marks awarded here). Strength: identified (large amounts of data can be gathered), explained/justified (makes the findings in relation to MPD more valid) and elaborated through appropriate contextualisation from the chosen study (information from all three personalities and family members). (3 marks awarded here). The response was therefore awarded 2+3=5 marks.

(c) Strength: 1 mark for identifying an appropriate strength (lots of rich, detailed data straight from the participant), 1 mark for the contextualisation (lots of information about Eve's home life and symptoms) but as there is no explanation/justification of the strength the response only gained 2 marks. Weakness: 1 mark for identifying an appropriate weakness (participants may lie). 1 mark for the contextualisation (Eve Black was known liar so could have given false information) but as there is no explanation/justification of the strength the response only gained 2 marks. The response was awarded 2+2=4 marks.

(d) Strength: 1 mark awarded for an appropriate strength (can gain both qualitative and quantitative data) + 1 mark for the qualitative evidence. As no quantitative evidence has been provided and no elaboration of the strength has been given, 2 marks were awarded. Weakness: No marks were awarded as this response refers to observation. The response was therefore awarded 2+0=2 marks.

### How the answer could be improved

A strength of self-reports is that they can generate large amounts of quantitative and qualitative data to give a comprehensive overview of the participant. Here Thigpen and Cleckley were able to gather quantitative data about the three different personalities, Eve White, Eve Black and Jane. For example they found that Eve White was quiet, demure and colourless, Eve Black was mischievous, childish and vain whereas Jane was interesting, sensible and mature; and qualitative data in relation to such things as IQ and brain activity. For example, both Eve White and Jane had EEG readings of 10½ - 11½ cycles per second whereas Eve Black's reading was 12-13 cycles per second.

A weakness of self-reports is that participants may respond with demand characteristics and give responses that they think the researcher is looking for. For example, as Eve Black was known to be a liar, she may have made up certain stories about what she did to fit in with Thigpen and Cleckley's expectations. This will affect the validity of the findings as they may not be true.

### General performance on the question

Many candidates were able to describe, explain/justify and support with evidence both an appropriate strength and an appropriate weakness of the self-report method. Some candidates however failed to support their answers with adequate evidence and some gave study-specific answers, which was not required.

## SECTION B, IN RELATION TO THIGPEN AND CLECKLEY – MULTIPLE PERSONALITY DISORDER, QUESTION 16(d)

### Question 16(d)

Outline the findings of your chosen study.

[8]

## SAMPLE ANSWERS

(a) Eve White had an IQ of 110 and a memory which was superior to her IQ level. Eve Black had an IQ of 104 and a memory which was on par with her IQ level, and inferior to Eve White's memory. Eve White's husband reported that she had been buying clothes out of character and Eve White could not remember doing it. Eve Black admitted that she did it. Eve Black did not care for Eve White. At the beginning, Eve Black was aware of Eve White and would talk about her but Eve White was not aware of Eve Black. Eve White on the Rorschach inkblot test tended to repression whereas Eve Black tended to regression. When Jane was revealed Thigpen and Cleckley decided she was the best personality as she was outgoing like Eve Black but sensible like Eve White, but they realised it was not their responsibility to decide which personality was best. Eve White was sensible, neat, colourless, serious. Eve Black was outgoing, mischievous, immature. On the EEG Eve White and Jane had very similar readings whereas Eve Black's was faster, bordering on psychotic.

**Mark** – Top band, 7-8 marks

(b) Eve White had an IQ of 110 whereas Eve Black had an IQ of 104. Eve White had a good memory which made it superior to Eve Black's. The Rorschach inkblot test showed Eve White was repressive and Eve Black regressive. Eve White was anxious, neat and colourless whereas Eve Black was vain, childish and devious. Jane however was interesting, calm and sensible.

**Mark** – Mid band, 4-6 marks

(c) Thigpen and Cleckley used various psychometric and physiological tests on Eve White and Eve Black. They conducted an EEG test where Eve White was found to have 11 cycles per minute and Eve Black 12.5 cycles per minute. They also used the Rorschach inkblot test as a projective test and Eve White was found to have a repressive personality type and Eve Black a regressive personality type. They also sent off Eve White and Eve Black's handwriting for analysis and both were declared to be by the same person.

**Mark** – Mid band, 4-6 marks

(d) Thigpen and Cleckley found Eve White was more intelligent and had a better memory than Eve Black. Three personalities were discovered altogether – Eve Black, Eve White and Jane.

**Mark** – Bottom band, 1-3 marks

## COMMENTARY

(a) The description of the results is accurate and detailed. Many fine details have been included e.g. IQ scores, memory test results and findings of the Rorschach test. Understanding, expression and use of psychological terminology are all good. Due to these factors, the answer was placed in the top band.

(b) Although rather lacking as far as 'description' is concerned this answer is accurate and includes several fine details e.g. IQ scores, memory test results and findings of the Rorschach test. Understanding is evident and the answer has some structure and organisation. It was therefore placed in the middle band.

(c) The results of the EEG test are inaccurate as the candidate has written 'per minute' instead of 'per second', but the difference between the two personalities is clear. The description is therefore mainly accurate but there are many omissions and more information should have been provided considering the level and time allowed. The response was adequate enough to be placed in the middle band.

(d) This was placed in the bottom band because the description of the results is very basic and no fine details have been included. As a result, the response was placed in the bottom band.

### How the answer could be improved

Not long into therapy Eve White's husband reported that Eve had gone into town and had bought an array of expensive clothes that were completely out of character. On her next visit to Thigpen and Cleckley she explained for the first time that she heard voices which were becoming more and more frequent and expressed a fear that she was going mad. This prompted the appearance of the 2nd personality – Eve Black. After this, the headaches and blackouts improved. However after 8 months of therapy the situation changed for the worse. Eve White's headaches and blackouts returned. During one session the 3rd personality – Jane appeared, indicating there were three personalities – Eve White, Eve Black and Jane. Eve White was found to be quiet, demure and colourless, Eve Black mischievous, childish and vain whereas Jane was found to be interesting, sensible and mature. In the psychological tests Eve White was found to have an IQ of 110 and a memory superior to her IQ. On the other hand Eve Black was found to have an IQ of 104 with a memory on par with her IQ. The Rorschach inkblot test showed that Eve White was emotionally repressed whilst Eve Black had a tendency to regress. EEG readings on the three personalities showed both Eve White and Jane to have of readings of 10½ - 11½ cycles per second whereas Eve Black's reading was 12 13 cycles per second.

### General performance on the question

This question saw the full range of marks being awarded. Some answers were very muddled and 'waffly', although there was evidence that many candidates knew this study really well.

## SECTION B, IN RELATION TO THIGPEN AND CLECKLEY – MULTIPLE PERSONALITY DISORDER, QUESTION 16(e)

### Question 16(e)

Suggest improvements to your chosen study.

[8]

## SAMPLE ANSWERS

(a) One way to improve the study by Thigpen and Cleckley is to use a larger sample, perhaps of 10 people who have MPD or at least claim to do so. This is because trying to generalise findings from one person, as in Thigpen and Cleckley's study is impossible, a larger sample would be more representative of MPD sufferers. These could be gathered through referral from psychiatric clinics, once doctors had asked the MPD sufferers' permission. As well as this, another way you could improve the study is by having an independent observer present who watches through a one-way mirror to make sure he does not become too involved with the study and remains objective, unlike Thigpen and Cleckley who appeared to show researcher bias particularly towards Eve Black who they found very attractive. Furthermore you could improve the study by asking all 3 personalities – Eve White, Eve Black and Jane – to give consent to taking part in the study as it may have been only Eve White who actually consented in the original study. In addition to this, another way you could improve the study is by repeating the psychometric tests such as the IQ test and inkblot tests on all 3 personalities more than once to make sure Eve White was not lying about her personality disorder. Finally you could improve the study by conducting it at Eve White's house rather than a clinic as this would be a more natural environment.

**Mark** – Top band, 7-8 marks

(b) One improvement would be to the sample, as investigating only one MPD patient of a 25 year-old woman called Eve White means results would not be generalisable to the wider MPD population. I would gather 20 patients suffering the same blinding headaches and gaps in memory symptoms, 10 male, 10 female, from 5 countries around the world (4 MPD participants from Tokyo, Dublin, London, Spain and Australia).

Another improvement would be to the method as a case study using self-report in Thigpen and Cleckley's study meant that Eve White knew she was being studied so may have played up to the researchers. I would place cameras in the MPD patients' homes to observe consistently and alterations on participants' behaviours in their own home to give more accurate and real life representations of changes in behaviour of people with MPD. They would be recorded 24 hours a day for a 2 week period to ensure all alterations in behaviour were recorded and these could then be analysed by Thigpen and Cleckley afterwards.

**Mark** – Third band, 5-6 marks

(c) Eve White could have been recorded at all times when she was at home, with the use of a hidden camera. A voice recorder could also have been given to Eve to record all utterances of each personality. There could also have been a number of researchers recording and observing Eve. Therapy sessions could also have been done at home. Qualitative data could have also been collected further by asking each personality to explain how they were feeling and thinking

**Mark** – Second band, 3-4 marks

(d) I would improve the study by conducting it in the participant's own home. I would increase the sample size and have an equal number of males and females and make sure participants came from more than one country. I would improve the procedure by having more researchers involved in gathering the data.

**Mark** – Bottom band, 3-4 marks

## COMMENTARY

(a) Several appropriate improvements have been suggested and justified. The description is detailed and understanding is evident. The improvements would be effective and have been linked to the chosen study. Implementation has been considered and this is appropriate to the level and time allowed. The response was therefore placed in the top band.

(b) Two appropriate improvements have been suggested. (It's a shame the candidate does not differentiate between a country and a city!) However, although both improvements have been justified and some thought has been given as to how they could be implemented, this is a basic answer which does not show any real 'creativity' or imagination specific to the chosen study i.e. it is the standard 'what I learned earlier' response. The answer was therefore placed in the third band.

(c) Several improvements have been suggested but none have been justified and little real thought has been given as to how they could be implemented. The answer is basic and lacks detail, although some understanding is evident. The response was therefore placed in the second band.

(d) The description is extremely basic and lacks any detail. No thought has been given as to how any of the suggestions could be implemented or why the suggestions are improvements. There is no contextualisation, so the response could apply to any of the named studies. The answer was placed in the bottom band.

### How the answer could be improved

One improvement to Thigpen and Cleckley's study would be to use a larger and more varied sample, perhaps of 10 people (5 males and 5 females, from different countries around the world e.g. one male and one female from each of the USA, France, South Africa, India and Japan) who have been diagnosed with MPD. This is because trying to generalise findings from one person, as in this study is impossible. These participants could be gathered through referral from psychiatric clinics and hospitals, once doctors had asked the patient's permission. Another improvement could be to move the study from the researchers' clinic to the participant's natural environment. This could be done by placing CCTV cameras and voice recorders at strategic places around the participant's home and / or workplace. This way behaviour can be continuously recorded. In addition to this, another way you could improve the study is by conducting all the psychometric tests such as the IQ test, memory test and inkblot tests on all 3 personalities as Jane did not attempt any of these tests. I would also conduct the tests more than once to make sure Eve White was not lying about her personality disorder.

### General performance on the question

A good range of marks were awarded for this question. Many candidates were able to suggest and justify two or more appropriate improvements. Real understanding was shown by those candidates who both explained why their suggestions would improve their chosen study and described how they would implement their suggestions. There were some instances where answers lacked contextualisation e.g. I would improve the sample by having more participants, or I would improve the ecological validity by conducting the study in a natural environment.

## SECTION B, IN RELATION TO THIGPEN AND CLECKLEY – MULTIPLE PERSONALITY DISORDER, QUESTION 16(f)

### Question 16(f)

Evaluate the improvements you have suggested to your chosen study.

[8]

## SAMPLE ANSWERS

(a) By using a larger sample, Thigpen and Cleckley would possibly be able to generalise their findings to lots of other MPD sufferers, especially if any common features were discovered. However interviewing 10 people to the same depth and level as Eve White would be extremely time consuming and hard to do. An implication of having an independent observer watch the interviews through a one-way mirror is that this would reduce the likelihood the results are biased and he could check his findings with those of Thigpen and Cleckley to see if there was inter-rater reliability. As well as this, by asking all 3 personalities of Eve to give consent to taking part in the study, you are making the study more ethical. However there is the chance that one personality may not give consent and this would cause problems for the researcher. Also by repeating the psychometric tests you will be improving both the reliability and validity of the findings. Finally by carrying out the study in Eve White's home would make the study have higher ecological validity as Eve's personalities may come out more easily in a familiar setting. However this would mean that there is a lack of control over confounding variables such as there being other family members around who might influence when and how personalities appear.

**Mark** – Top band, 7-8 marks

(b) By getting another psychologist to analyse Eve to get a different perspective and opinion on the disorder may have allowed for more detail and qualitative data to be obtained. A 'blind' psychologist being used as part of the study would have also reduced extraneous variables concerning Thigpen and Cleckley twisting any data in order to get the results they required – a 'blind' experimenter would not manipulate data in this way. It may have caused further problems however concerning ethical issues with keeping information confidential within the therapy sessions. One problem with trying to both make the sample more representative and larger is that multiple personality is an extremely rare disorder so obtaining the extra participants may be hard, timely and cost-consuming. However if similar findings are found to that of Eve by studying more people with MPD it can make the findings more reliable and valid.

**Mark** – Third band, 5-6 marks

(c) By using a video tape to record interviews will improve accuracy and reliability. However Eve White may be distracted by the camera and the time spent in interviews doesn't provide as many useful findings because of her being distracted during the time. By using more people in the sample with the same physical condition as Eve, more of the population is represented so possible evidence for MPD can be generalised to all those who suffer from severe headaches followed by blackouts.

**Mark** – Second band, 3-4 marks

(d) Using a larger sample of 5 females and 5 males with MPD would make the findings more generalisable. In Thigpen and Cleckley's study the results of the psychometric tests and projective tests cannot be generalised further than the sole participants in the study: Eve White/Eve Black/Jane. However sending letters to multiple psychologists to get the sample has serious time and cost implications. Doing the study in England will also increase the generalisability of the results. This is because the participant will be from a different country to America so the results could be compared and made more generalisable, specifically to MPD sufferers in England.

**Mark** – Second band, 3-4 marks

## COMMENTARY

(a) The evaluation of the suggested improvements is appropriate and detailed. Good understanding is shown and the answer is well linked to the chosen study. The answer is competently structured and organised and is appropriate to the level and time allowed. The response was placed in the top band.

(b) Two improvements have been evaluated, although the first evaluation is rather muddled with, at times, a strange use of use of psychological terminology e.g. 'reduced extraneous variables' which may infer a lack of true understanding. Therefore, although evaluation is reasonable, expression is somewhat limited. As a result, the response was placed in the third band.

(c) Evaluation though appropriate is basic and lacks detail. Expression is limited and there is little evidence of any real understanding. Therefore, the response was placed in the second band.

(d) Although the evaluation of improvements is appropriate it is very basic and lacks detail. Some understanding is evident but the implications of the improvements could have been developed much further. The answer was placed in the second band.

### How the answer could be improved

The use of a larger sample would possibly mean the findings could be generalised to other MPD sufferers, especially if any common features were discovered. However finding 5 males and 5 females who suffer with MPD is likely to prove very difficult as it is a very rare illness. Interviewing 10 people from different countries to the same depth and level as Eve White would be extremely time-consuming and hard to do, especially as there are likely to be language difficulties and behaviour may be interpreted differently in different cultures. By using CCTV and voice recorders to record the behaviour of the different personalities will improve accuracy reliability and validity as Eve will be behaving naturally in an ecologically valid environment. However Eve White may become aware of the recording devices and then become distracted and so not behave naturally. If she is aware of being recorded she could respond to demand characteristics by making the various personalities appear more / less regularly and or social desirability with Eve Black 'behaving herself' and not showing her true behaviour. There are also ethical issues with recording Eve White in her own home and at work without her knowledge. She would not have given consent to being recorded, not knowing the devices were there means she would be deceived and if she became aware that the behaviour of her different personalities was being recorded she could become extremely distressed, especially if she was then shown the recordings of her different personalities when they were 'out'.

### General performance on the question

Many candidates, although able to evaluate their improvements failed to contextualise their answers to their chosen study e.g. by increasing my sample size and including both males and females I will get a more representative sample so I can generalise my findings to the whole population. Such a statement could apply to any of the three studies offered in this section. Furthermore, many evaluation points were extremely basic and failed to show any real understanding e.g. conducting the study in a natural environment would give a better representation of real life.

# SECTION B, IN RELATION TO REICHER AND HASLAM – BBC PRISON STUDY, QUESTION 16(a)

## Question 16(a)

Outline why your chosen study was conducted.

[2]

## SAMPLE ANSWERS

(a) To investigate the behaviour of groups with unequal power and resources and show that a social study can be conducted in an ethical way.

**Mark – 2 marks**

(b) Reicher and Haslam's was a prison study conducted to see how people act tyrannically in groups and to see which variables can lead to tyranny.

**Mark – 1 mark**

## COMMENTARY

(a) The candidate has given two clear reasons why this study was conducted and the response has been contextualised through the reference to the groups having unequal power and resources. The full 2 marks were therefore awarded.

(b) This was considered a partial/vague answer as the candidate needed to develop their response further e.g. either in relation to 'variables that can lead to tyranny', or by referring to how people act tyrannically when in groups of unequal power and privilege between them/the role of social, organisational and clinical factors in group behaviour. 1 mark was awarded because of this.

### How the answer could be improved

- To study the conditions that lead individuals to (i) identify with their group; (ii) accept or challenge intergroup inequalities.
- To carry out a prison study with ethical procedures that would ensure no harm to the participants.

### General performance on the question

This was a well answered. Good responses were clear and fully contextualised. Weak responses included careless errors e.g. suggesting that this study was to replicate Zimbardo's study. Those candidates who chose this study struggled to get a reasonable answer.

## SECTION B, IN RELATION TO REICHER AND HASLAM – BBC PRISON STUDY, QUESTION 16(b)

### Question 16(b)

Describe how the self-report method was used to gather data in your chosen study.

[4]

## SAMPLE ANSWERS

(a) The self-report method was used to gather data in the study from the use of daily psychometric tests. These provided the researchers with quantitative data. The psychometric tests were used to measure clinical and social variables. For example clinical measuring of how the guards and prisoners were feeling psychologically once they were in the prison – e.g. if the prisoners felt depressed, and social variables measuring how the prisoners and guards identified with their roles and whether they identified with their group or not.

**Mark** – 4 marks

(b) They administered psychometric tests in the form of questionnaires for the prisoners and guards to fill out. They did daily tests throughout the study for such things as group identification, cognitive alternatives, and depression using rating scales.

**Mark** – 3 marks

(c) Throughout the study participants were asked to visit a room secluded and separated from the prison environment to answer questions in an interview procedure with the experimenters in which they expressed their feelings about events ensuing in the study. Also, a thorough debriefing procedure was done by Reicher and Haslam after the aftermath of the 11 day study in which participants were questioned and interviewed..

**Mark** – 1 mark

(d) The self-report method was used through the daily video diaries that the participants were made to keep. They recorded their thoughts, feelings and opinions and expressed what they thought away from others. This collected qualitative data which was observed by Reicher and Haslam.

**Mark** – 0 marks

## COMMENTARY

(a) This was awarded the full 4 marks as the candidate has described in depth one way in which data was gathered using the self report method i.e. through daily psychometric testing throughout the study.

(b) This is a fairly accurate description of how the self-report method was used in Reicher and Haslam's study. The reference to testing for group identification is somewhat ambiguous and the variables measured could have been given in more detail. However, the response was considered worthy of being awarded 3 marks as it goes beyond being 'partial or vague'

(c) The first situation given here is not mentioned in the original study. However, 1 mark was awarded for the reference to the debriefing after the study was terminated.

(d) 'Video diaries' are not mentioned in the original study so as answers must be based upon the original study, this response was not awarded any marks.

### **How the answer could be improved**

- Self report data was gathered firstly in the three-phase screening process which reduced the initial pool of 332 applicants to 27 men. During this process potential participants completed a battery of psychometric tests that measured both social variables (e.g. authoritarianism) and clinical variables (e.g. anxiety). Self report data was also gathered through daily psychometric testing throughout the study in which participants had to apply rating scales to a variety of tests.
- Self report data was gathered firstly in the three-phase screening process during which process potential participants completed a battery of psychometric tests that measured both social variables (e.g. authoritarianism) and clinical variables (e.g. anxiety). Self report data was also gathered through daily psychometric testing throughout the study. Participants had to apply rating scales in a variety of tests to measure how social (e.g. awareness of cognitive alternatives), organisational (e.g. compliance with rules) and clinical (e.g. depression) variables developed throughout the duration of the study. Participants remained for a further day after the study was brought to a halt, to undertake a series of structured debriefings designed to obtain and provide feedback on their experience.

### **General performance on the question**

For this question, only the more able candidates appreciated that the self-report method was used to gather data in at least two ways in their chosen study. For example, psychometric tests during the screening process, psychometric tests throughout the study, during the debriefing at the end of the study, or were able to describe in depth one way the self-report method was used. Some candidates who focused on this study referred to going to special rooms to talk to the researchers/using video diaries. These suggestions were not deemed creditworthy as they are not referred to in the original study.

## SECTION B, IN RELATION TO REICHER AND HASLAM – BBC PRISON STUDY, QUESTION 16(c)

### Question 16(c)

Describe one strength and one weakness of gathering data using self-reports. Support your answer with evidence from your chosen study. [6]

## SAMPLE ANSWERS

(a) One strength of gathering data using self-reports is that it can provide the researcher with quantitative data. For example in the study by Reicher and Haslam the psychometric tests provided quantitative data which can be easily statistically analysed and compared. For example you can compare how the prisoners were feeling compared with the guards psychologically by statistically analysing their self-reports which used rating scales.

One weakness of gathering data using self-reports is that social desirability could affect the participants' answers – they may lie. For example in this study some of the guards could have lied in their psychometric tests and said they weren't identifying with their group as they wanted to appear socially acceptable – they didn't want to appear as though they were enforcing hierarchy in the prison as this could be seen as being too dominant. The findings therefore lack validity.

**Mark – 6 marks**

(b) One strength is you can gain insight into the participant's thoughts, feelings and opinions. For example in Reicher and Haslam they found out how depressed they were and if they had identified with their group. They could also see that after permeability of roles was stopped prisoners identified better with their group. They may not have known such information had they not used the self-report questionnaires.

A weakness would be there may have been social desirability. For example the prisoners may not have wanted other people to really know how they felt so could have changed their answers, affecting the validity of the findings.

**Mark – 6 marks**

(c) One strength of gathering data using self-reports is that much quantitative and qualitative data can be gained. For example in the BBC prison study the quantitative data gained was used to create graphs for comparison which are quite easy to analyse.

One weakness of self-reports is that they may become a laborious task, therefore participants become bored and circle any answer. For example, in the study by Reicher and Haslam, the participants had to complete daily questionnaires using a rating scale which means they could have become bored and answered using any random number. This means the findings are not a true representation of the prisoners and guards feelings so the results would lack validity.

**Mark – 4 marks**

(d) A strength is that data is in depth and reflective of the participant's true thoughts and feelings. For example, in Reicher and Haslam's study, interviews were largely unstructured as the interviewer was not in the same room as the participant and through this the experimenters collected extra data such as how the guards were trying to hide their empathy for the prisoners.

A weakness however is that self-report is time consuming, so when the participants were being interviewed the observation of their interaction with the other participants could not be done, reducing the strength of the findings and ecological validity as a result.

**Mark – 2 marks**

## COMMENTARY

(a) Strength: identified (quantitative data can be gathered), explained (allows easy statistical analysis and comparison to be made) and elaborated appropriately through evidence from Reicher and Haslam's study. Therefore the strength was awarded 3 marks.

Weakness: identified (social desirability), explained (means findings lack validity) and elaborated appropriately through reference to Reicher and Haslam's study. 3 marks awarded. The response therefore gained 3+3=6 marks.

(b) Strength: identified (insight into thoughts and feelings), explained (without the use of questionnaires this information could not have been gathered) and elaborated through appropriate evidence. 3 marks

Weakness: identified (social desirability), explained (results may not be valid) and elaborated through appropriate evidence. 3 marks. Overall the response was therefore awarded 3+3=6 marks.

(c) Strength: identified (quantitative and qualitative data can be gathered), but as there is no explanation or appropriate evidence the response could only gain 1 mark.

Weakness: identified (questionnaires may be laborious and so subject to the boredom effect), explained/justified (the findings will lack validity as any answer is provided) and supported with appropriate evidence from the chosen study. 3 marks were awarded here. Overall the candidate therefore gained 1+3=4 marks.

(d) 1 mark for the identification of an appropriate strength (can gain information on an individual's true thoughts and feelings). This however has not been elaborated or supported by appropriate evidence. Weakness: 1 mark for the identification of an appropriate weakness (interviews can be time-consuming). This however has not been elaborated or supported by appropriate evidence. The answer gained 1+1=2 marks.

### How the answer could be improved

A strength of self-reports is that they can gain insight into the participant's thoughts, feelings, opinions and mental states. Here Reicher and Haslam, through the use of the daily psychometric tests, were able to measure how social (e.g. awareness of cognitive alternatives), organisational (e.g. compliance with rules) and clinical (e.g. depression) variables developed in both the guards and prisoners throughout the duration of the study. They may not have known such information had they used another method such as observation.

A weakness of self-reports is that responses may have a social desirability bias because participants give an answer which they think makes them look good or is acceptable to society. For example, some of the guards could have lied in their psychometric tests and said they weren't identifying with their group as they wanted to appear socially acceptable – they didn't want to appear as though they were mean and nasty by giving out too many punishments as this could be seen as being too dominant and cruel. The findings therefore lack validity.

### General performance on the question

Many candidates were able to describe, explain/justify and support with evidence both an appropriate strength and an appropriate weakness of the self-report method. Some candidates however failed to support their answers with adequate evidence, while others gave study-specific answers, this was not appropriate for this question.

## SECTION B, IN RELATION TO REICHER AND HASLAM – BBC PRISON STUDY, QUESTION 16(d)

### Question 16(d)

Outline the findings of your chosen study.

[8]

## SAMPLE ANSWERS

(a) Initially the prisoners acted individually as they wanted promotion to guard, but after impermeability of roles on day three they formed a stronger group identity and worked against the guards. For example when they stole the guards' keys. The guards failed to form a strong group identity as they couldn't agree on their priorities and were wary of using their authority. Their authoritarianism was higher at the beginning but went down after day 3 once the roles were set. The introduction of the trade unionist on day 5 was not needed to suggest cognitive alternatives though he was able to make them think more about the situation they were in and how they could change things. Compliance with the guards fell down further on day 6 when prisoners took over the guards' quarters and drew up plans for a new commune, but this broke down within a day. Depression for the guards at the end was higher than for the prisoners though in the beginning it was the other way round.

**Mark** – Top band, 7-8 marks

(b) The results of Reicher and Haslam's study showed that up until day 3 whilst the roles were permeable (one prisoner had the opportunity to be promoted to a guard), that the prisoners appeared segregated and did not join together as a group. After day 3 when the promotion was fixed and the roles were impermeable, the prisoners joined together to start to humiliate the guards. On day 5 when a trade unionist was introduced it was shown that he did bring about a new way of thinking – a cognitive alternative as a forum was introduced where once a day the prisoners and guards would meet to discuss their thoughts. However when the trade unionist left the commune broke down and there was again conflict between the prisoners and the guards. Another finding from the study showed that at first the guards didn't identify with their role as they gave away resources to the prisoners – for example, food. However when a new authoritarian regime was introduced, the guards did identify with their role as they asked for sunglasses and black berets.

**Mark** – Top band, 7-8 marks

(c) As time went on during the study the prisoners' group identification increased, however the guards' group identification decreased. Throughout the study the prisoners' self-efficacy increased, but the guards' self-efficacy decreased. The depression levels of the prisoners decreased throughout the study, but the depression levels of the guards increased. When the first prisoner was promoted to guard, the prisoners then started to behave how they thought would get them promoted too.

**Mark** – Mid band, 4-6 marks

(d) In Reicher and Haslam's study, they found that the guards did not identify with their roles of authority and so were unable to operate as a team and form some form of government. They also found that when the prisoners were made aware of potential movement of role from prisoner to guard that two prisoners consciously wanted to become guards and expressed their reasons by saying, "They get all the luxuries". Furthermore they found that once the permeability of roles was removed the prisoners identified with their roles and appeared to be able to work as a group better. They also found that the desire to create a commune united them all but once this commune was existent a few participants decided to wreck havoc and cause tyranny.

**Mark** – Mid band, 4-6 marks

## COMMENTARY

(a) There is plenty of accurate information here and several fine details have been included e.g. the breaking down of the commune within a day and the reference to depression levels reversing as the study progressed. Understanding and use of appropriate psychological terminology are good. Detail is appropriate for the level and time allowed. For these reasons the answer was placed in the top band.

(b) Although there are some inaccuracies e.g. the trade unionist did not need to introduce cognitive alternatives per se, there are plenty of accurate details so the response was considered worthy of being placed in the top band.

(c) Although the description is basic and lacks detail, what is written is mainly accurate with some understanding. The last statement however is incorrect. Therefore, the answer was placed in the second band.

(d) The description of results, though sparse, is mainly accurate and some fine details have been included e.g. the statement by one of the prisoners expressing why they wanted to become a guard. Some understanding is evident but much more could have been included when one considers the level and time allowed. The response was therefore placed in the second band.

### How the answer could be improved

Overall, the guards failed to internalise their role and failed to develop a group identity. Initially most of the prisoners acted individually as they wanted to be promoted to the role of guard. Two prisoners (JE and KM) made particular efforts to be promoted. However once group impermeability was introduced, the prisoners began to develop a much stronger sense of shared identity and to develop more consensual norms. There was a move from compliance to conflict with the guards e.g. to see how the guards would respond, prisoner JE threw his lunch plate to the ground and demanded better food. They also began to envisage changes to the existing hierarchy and believe they could achieve them. The introduction of the new prisoner on Day 5 (a trade union official) was not needed to suggest cognitive alternatives as they had already surfaced, rather he was able to suggest additional alternatives to the status quo. On Day 6, dissent came to a head and some prisoners broke out of their cell and occupied the guards' quarters. The guards' regime therefore became unworkable. Terms for a new commune were drawn up but within a day this was in crisis because two ex-prisoners broke communal rules. A further harsher prisoner-guard regime was proposed but for ethical reasons could not be implemented so the study was stopped.

### General performance on the question

This question saw the full range of marks being awarded. Some answers were very muddled and 'waffly' and there was little evidence that candidates knew this study really well.

## SECTION B, IN RELATION TO REICHER AND HASLAM – BBC PRISON STUDY, QUESTION 16(e)

### Question 16(e)

Suggest improvements to your chosen study.

[8]

## SAMPLE ANSWERS

(a) The study could be improved by increasing the sample size from 15 participants to 50 participants. This could be done through using the screening process already used by Reicher and Haslam and building a bigger prison environment to accommodate more than 15 participants. This would consequently improve the population validity.

Additionally, the ecological validity of the study could be increased by having the study take place in a real prison with real guards. This could be done by using a spare wing of a real prison and asking the staff if they would take part in the study.

**Mark** – Third band, 5-6 marks

(b) An improvement I would make to the study is to have it take place at an actual prison as opposed to a film studio, or a recreation of a prison. I would ensure that the prison is no longer in use but is a safe environment.

I would also include women in the study by having an equal number of men and women in the study. They too would undergo the same psychometric tests and screening as the original sample.

I would also increase the number of participants from 15 to 50 by not reducing the initial pool of participants by as much to recreate a small prison rather to have just a few prisoners and guards.

**Mark** – Third band, 5-6 marks

(c) The study could be improved by conducting it in a real prison with real prisoners to covertly observe a group of new prisoners to see how they react to prison life. This will test social identity theory as they may or may not identify with the existing prisoners. It is also more ecologically valid because it will be taking place in a real prison environment with real events, rather than a prisoner being promoted to a guard. Furthermore, demand characteristics are less likely as they will not know they are being observed. However, it may be unethical as consent from the prisoners cannot be gained, only from the prison.

Another improvement could be to conduct the study again using female prisoners and then comparing the results to see if there are gender differences in SIT which would make the study more generalisable.

**Mark** – Second band, 3-4 marks

(d) One improvement that could be made to the Reicher and Haslam study would be to not show the study on the BBC (television). Instead of showing the prison study on TV I would conduct the study where it is not shown on TV, for example I would still conduct the study in a field setting where it represents a real life prison setting however it would not be set up for a TV show.

A second improvement that I would make to the study would be to use CCTV cameras. For example I would have 2 CCTV cameras in each room and in each of the prison cells and this would act as a recording of observational data. Another improvement that I would make to this study would be to change the sample. For example I would have 15 males and 15 females, making a total sample of 30 participants. I would use the same method of sample selection that Reicher and Haslam did. I would recruit my 30 participants through a newspaper advert and a direct mailing advert. Another improvement that I would make to this study would be to carry out the study in a real live prison – for example in the London Prison where there are real prisoners and real guards.

**Mark** – Second band, 3-4 marks

## COMMENTARY

(a) Two appropriate improvements have been suggested and justified. Some thought has been given as to how these improvements might be implemented. Both improvements could however have been developed further. The response was placed in the third band.

(b) The candidate has suggested several appropriate, though basic improvements and the description is reasonable. Some consideration has been given to how the improvements could be implemented but these could have been developed much further. The response was placed in the third band.

(c) Although some appropriate improvements have been suggested, the first paragraph is very muddled and much of its content is evaluative and so should have been included in 16f. The description is basic and lacks detail though some understanding is evident. The response was placed in the second band.

(d) Although several appropriate improvements have been suggested, no justifications or explanations have been offered for the improvements. The description is basic and lacks detail though some understanding is evident. The response was placed in the second band.

### How the answer could be improved

One improvement to Reicher and Haslam's study could be to increase the sample size. Reicher and Haslam only had 15 participants which once they were divided into 5 guards and 10 prisoners was not very representative of a real prison environment. This could be done through using the screening process already used by Reicher and Haslam but reducing the initial pool to 40. I could then match them into fours on selected personality variables and then randomly allocate them three prisoners to one guard so I would have 30 prisoners and 10 guards.

A second improvement I could make would be to conduct the study again using females as Reicher and Haslam only used males. Females could be selected and allocated to the roles of guards and prisoners in the same way as in the original study.

A further improvement that I could make to the study would be to use CCTV cameras, instead of observation by the researchers to gather data. For example I could have CCTV cameras in every room of the simulated prison e.g. each cell, the guards' observation post, the exercise area, and the main atrium. These cameras would act as a recording of observational data.

### General performance on the question

A good range of marks was awarded for this question. Many candidates were able to suggest and justify two or more appropriate improvements. Real understanding was shown by those candidates who both explained why their suggestions would improve their chosen study and described how they would implement their suggestions. There were some instances where answers lacked contextualisation e.g. I would improve the sample by having more participants and I would improve the ecological validity by conducting the study in a natural environment.

## SECTION B, IN RELATION TO REICHER AND HASLAM – BBC PRISON STUDY, QUESTION 16(f)

### Question 16(f)

16(f). Evaluate the improvements you have suggested to your chosen study.

[8]

## SAMPLE ANSWERS

(a) By changing the sample size from 15 to 50 participants the population validity will increase and the data gathered will be more generalisable to a wider target population as there are more people with varying backgrounds in the study and so more representative of the general population. However recruiting an additional 35 participants would mean creating a new prison environment which would be very time and money consuming.

By changing the artificial prison environment to a real prison environment with real guards, the ecological validity of the study would be increased. If ecological validity increases the data gathered will be more generalisable to real world group behaviour. Although making this change would be near impossible and resource consuming. Doing this change would be very impractical as it may involve the movement of real prisoners to another prison which may not be possible.

**Mark** – Third band, 5-6 marks

(b) By not showing the study on television it would mean reducing demand characteristics that the prisoners and guards may show as they knew they were being watched. This would improve the validity of the study as it's measuring what's been intended to be measured – the natural behaviours of the prisoners and guards in a prison setting. The second improvement to the study of not having real life observers but just CCTV cameras would also reduce demand characteristics. However a weakness of just having CCTV cameras to obtain observational data is that the researchers could miss behaviours that were not recorded on the film. By changing and improving the sample from 15 males to 15 males and 15 females (30 participants), it makes the study into prison life more generalisable. However it could be hard to find 15 females who are willing to put up with the solitude and hunger in a prison study.

**Mark** – Third band, 5-6 marks

(c) One strength of using a real prison and real prisoners is that it will make the study more ecologically valid because it will not be taking place in an artificial environment, as Reicher and Haslam's study originally was. Another strength is that demand characteristics are reduced because the participants will not know they are being filmed so they will act naturally. A weakness of this improvement is that it will make the study less ethical because the participants themselves will not be able to give consent, the right to withdraw and cannot be debriefed as they were in the original study.

**Mark** – Second band, 3-4 marks

(d) Using females in my study would increase the generalisability of my results to the whole population. However it may be hard to find an equal number of males and females willing to participate. Carrying out the study in a natural environment would increase the ecological validity of my results as behaviour would be more naturalistic. However it may be difficult to have access to observe behaviour in a real prison. By using a larger sample size I would be increasing the generalisability of the results as participants with an even wider range of ages and occupations would be selected. However it would be more difficult to carry out and observe 50 participants.

**Mark** – Bottom band, 1-2 marks

## COMMENTARY

(a) Two improvements have been appropriately evaluated but although understanding is evident the evaluation is basic and lacks any real depth and the links to the chosen study could have been developed further. However both improvements have been evaluated for both positive and negative features and so the response was placed in the third band.

(b) Although the candidate has written a lot, many points have not been fully developed and expression is limited. Some points are incorrect e.g. the aim was not to see natural behaviours of prisoners and guards in a prison setting, and the response lacks detail. It was therefore placed in the third band.

(c) This evaluation is basic and lacks any real detail. There is some understanding but expression is limited and the answer is only loosely linked to the chosen study. The response was placed in the second band.

(d) Although several appropriate evaluative points have been made, the answer lacks depth and evidence of real understanding. The answer is barely linked to the chosen study so was placed in the bottom band.

### How the answer could be improved

By changing the sample size from 15 to 40 participants the population validity will increase and the data gathered in relation to studying the conditions that lead individuals to identify with their group and accept or challenge intergroup inequalities will be more generalisable to a wider population as the sample should be more representative of the general population. A difficulty of doing this would be that a larger sample would mean that a larger prison environment would have to be built so more space would be needed. Elstree television studios might not be big enough and so another suitable place would have to be found. This would be very time and money consuming. Conducting the study again using females would mean that results from both studies could be compared to see if there are gender differences in SIT. If there are similarities the findings will be generalisable to both genders but if there are differences further research can be conducted into the topic to see if reasons can be found for these. It could however be hard to find 15 females who are willing to put up with the solitude and hunger in a prison study. A problem of using CCTV cameras to gather observational data instead of the researchers themselves is that some behaviours could be missed if the cameras are not able to cover every area of the prison environment. Although the use of CCTV cameras in a real prison is not an ethical issue as they are in permanent use in modern prisons, if the participants in this study, who would not be real guards and prisoners, are not aware they are being filmed they are being deceived. They may also feel their privacy is being invaded and if they do become aware of being filmed they may become distressed. A strength however of using CCTV cameras is that even if participants know their behaviour is being recorded, because they are so common in today's society, they are likely, over time, to forget the cameras are there and play their roles as guards and prisoners in a natural way. This means the data gathered will be valid.

### General performance on the question

Many candidates, although able to evaluate their improvements failed to contextualise their answers to their chosen study e.g. by increasing my sample size and including both males and females I will get a more representative sample so I can generalise my findings to the whole population. Such a statement could apply to any of the three studies offered in this section. Furthermore, many evaluation points were extremely basic and failed to show any real understanding e.g. by conducting the study in a real prison would give a better representation of real life.

# SECTION C, IN RELATION TO THE BEHAVIOURIST PERSPECTIVE

## QUESTION 17(a)

### Question 17(a)

Briefly outline the behaviourist perspective.

[2]

## SAMPLE ANSWERS

(a) The behaviourist perspective states that we are born as blank slates and that all our behaviour is learned from our experiences and the environment around us after birth.

**Mark – 2 marks**

(b) The behaviourist perspective states that our behaviour is formed and affected by the rewards and punishments given by society.

**Mark – 1 mark**

(c) When children grow and mature they are able to look at things differently. Not because of better language or more intelligence but because of the ability to look at things in a more abstract way.

**Mark – 0 marks**

## COMMENTARY

(a) This is a clear outline that refers to both the perspective and behaviour. It was awarded the full 2 marks.

(b) This is a rather vague response that does not have a clear link to the perspective and makes no reference to the fact that behaviour is learned. The response was therefore only awarded 1 mark.

(c) There is no link to either the perspective or behaviour and this response really applies to the developmental approach so this response was deemed not creditworthy and so gained 0 marks.

### How the answer could be improved

- The behaviourist perspective holds that when we are born our mind is a tabula rasa (blank slate) and that all behaviour is learnt from the environment after birth.
- The behaviourist perspective sees the majority of all behaviour as being learned from the environment after birth.

### General performance on the question

This was generally well answered. Some candidates failed to refer to the fact that this perspective holds that behaviour is learned and therefore gave answers that related more to the social approach than the behaviourist perspective.

# SECTION C, IN RELATION TO THE BEHAVIOURIST PERSPECTIVE

## QUESTION 17(b)

### Question 17(b)

Describe how the behaviourist perspective could explain why some children become aggressive.

[4]

## SAMPLE ANSWERS

(a) The behaviourist perspective suggests that behaviour is a learnt response and is affected / influenced by the environment we are in. Aggression is a learnt response affected by whether children are exposed to aggressive behaviour displayed by an adult. For example, it was evidenced in Bandura's study on aggression that when children observed an aggressive model they later imitated this behaviour and displayed many more aggressive acts compared to children who were exposed to either a non-aggressive model or no model at all. This suggests that aggression is a response that children learn from observing how others around them behave.

**Mark** – Top band, 3-4 marks

(b) The behaviourist perspective could explain why some children become aggressive through the social learning theory. This is when our behaviour is shaped through our observations of our surrounding environment. For example, Bandura's study showed that when a child observed an adult model behaving aggressively, this child was more likely to imitate this aggression than if they had observed a non-aggressive model. This shows how the social learning theory is applicable in behaviourism – we can learn aggressive behaviour through the observation of aggressive models in our environment.

**Mark** – Top band, 3-4 marks

(c) The behaviourist perspective could explain why children become aggressive by looking at how children observe other people's behaviour, imitate it and after behaving like that in one situation do it again in another. If they see aggressive people they are likely to imitate the same behaviour later. When the model in Bandura's study was aggressive the children were also aggressive against the bobo doll.

**Mark** – 2 marks

## COMMENTARY

(a) The description is detailed and accurate, showing good understanding. There are clear links to the perspective and the elaboration in relation the Bandura's study is appropriate. The answer was placed in the second / top band.

(b) The description is detailed and accurate, showing good understanding. There are clear links to the perspective and the elaboration in relation the Bandura's study is appropriate. The answer was placed in the second / top band.

(c) 2 marks (bottom band) were awarded here as there is no clear link to the behaviourist perspective and the supporting evidence from Bandura is very weak.

**How the answer could be improved**

The behaviourist perspective holds that the majority of behaviour is learned and not innate. This perspective can therefore explain why individuals are aggressive since, according to Bandura, people learn behaviour through specific social learning processes: they observe significant others modelling behaviours which, when the appropriate opportunity arises, they copy or imitate. This was shown in his 'Bashing Bobo' study in which children observed a model acting aggressively towards a bobo doll. Later, when in a similar situation, even though the original model was no longer present, the children imitated many of the model's physical and verbal acts of aggression, showing aggressive behaviour had been learned.

**General performance on the question**

Overall, this question was better answered than in previous sessions. Many candidates were able to describe appropriate evidence (Bandura) or another appropriate example to show how aggression could be learned. Links to the perspective were better than in previous sessions with good candidates making appropriate links to either the social learning theory or operant conditioning.

# SECTION C, IN RELATION TO THE BEHAVIOURIST PERSPECTIVE

## QUESTION 17(c)

### Question 17(c)

Describe one similarity and one difference between any experiments that can be viewed from the behaviourist perspective. [6]

## SAMPLE ANSWERS

(a) One similarity between the studies by Bandura and Skinner in the behaviourist perspective is that they both use highly controlled settings. For example in Skinner's study he used a 'Skinner box' which was standardised and controlled for all of the animals when looking at operant conditioning in rats and pigeons. Likewise Bandura's study was controlled because all the children in the second stage were allowed to play with the same toys, for the same amount of time before being told they could not play with them any longer.

One difference between the studies between Bandura and Skinner was the type of participants used. For example in Skinner's studies into operant conditioning he used numerous animal participants – rates and pigeons. However Bandura used 72 children from Stanford University Nursery.

**Mark – 6 marks**

(b) One similarity between Bandura's study and Piliavin's study is that they both contain unrepresentative sample. Bandura's study used children from Stanford University Nursery. This means they may not be representative of children from other areas of the world. Piliavin's sample was an opportunity sample of those on the 8th Avenue New York subway between 11am-3pm. This means that those who work during the day are not included in the study so results can only generalised to those who use the subway between the stated times.

One difference between the studies is that Bandura's experiment was conducted in a laboratory, whereas Piliavin's was conducted in natural conditions. Bandura's study took place in a laboratory at Stanford University and Piliavin's study was conducted on the 8th Avenue IND line in New York.

**Mark – 5 marks**

(c) One similarity between Reicher and Haslam's BBC prison study and Bandura's study of aggression is that both of them involve setting up an artificial situation with the independent variable being altered by the experimenter and the participant's behaviour was recorded (so they were both lab experiments).

A difference is that Bandura's study of aggression is studying how children are affected by the environment while Reicher and Haslam's BBC prison study is studying how adults are affected by the environment.

**Mark – 2 marks**

(d) Freud and Samuel and Bryant both believed children go through different stages as they grow up. Samuel and Bryant believed in the cognitive development and that a child's ability to conserve gets better as they go through the stages. Freud also believed children go through different stages but in his study this was linked to the Oedipus complex.

A difference is that Freud used a case study with Little Hans as the only participant and he got information about what Hans did in his home and natural environment – it is high in ecological validity. Samuel and Bryant used a controlled lab experiment in their study. This may have affected the results and it was low in ecological validity.

**Mark – 0 marks**

## COMMENTARY

(a) Similarity: identified (highly controlled settings) and supported by appropriate evidence from the two named studies.

Difference: identified (different types of participant) supported by appropriate evidence from two appropriate studies.

The response was awarded 3+3=6 marks.

(b) Similarity: Identified (unrepresentative samples) and supported by appropriate evidence from both Bandura and Piliavin. 3 marks.

Difference: identified (Bandura – laboratory experiment, Piliavin in natural conditions i.e. field experiment) and supported with appropriate evidence from Piliavin. The evidence from Bandura is too vague to be creditworthy – in fact part of the experiment was conducted in the nursery which was set up to simulate laboratory conditions. 2 marks. The response therefore gained 3+2=5 marks.

(c) Similarity: merely identified (artificial situation/manipulated IV) with no evidence from either of the named studies.

Difference: merely identified (Bandura used children whilst Reicher and Haslam used adults – to study the effects of the environment) with no evidence from either of the named studies.

The response was therefore awarded 1+1=2 marks.

(d) Similarity + Difference: the candidate is referring to inappropriate studies so 0 marks were awarded.

### How the answer could be improved

One similarity between two experiments that can be viewed from the behaviourist perspective is that both the experiment conducted by Bandura and the experiment conducted by Reicher and Haslam lacked ecological validity. Bandura's experiment involved an adult model coming into a room where a child had been sat at a table to play with potato prints and picture stickers. The model went to an opposite corner of the room and either acted aggressively towards a bobo doll or ignored the bobo doll and played in a subdued manner with a tinker toy set. This did not reflect a real-life situation. Likewise, Reicher and Haslam's experiment was not only conducted in a specially constructed mock prison at Elstree Film Studios in London but also the participants who were randomly assigned to the roles of prisoners and guards were not genuine prisoners or guards so again this did not reflect real life.

One difference between two studies that can be viewed from the behaviourist perspective is the sampling method used: Reicher and Haslam used the volunteer sampling method by recruiting participants through advertisements in the national press and through leaflets whereas Piliavin used the opportunity sampling method of whoever travelled on the 8th Avenue New York subway between 11am - 3pm from April 15 to June 26, 1968.

### General performance on the question

Many candidates failed to respond to the question properly and answered in relation to the behaviourist perspective rather than experiments that can be viewed from the behaviourist perspective. Furthermore, many candidates compared inappropriate studies because they did not seem to be fully aware of the features of an experiment. Milgram is generally considered a controlled observation conducted under laboratory conditions. An experiment is defined as follows: 'an experiment involves the manipulation of an independent variable in order to see its effect on a dependent variable' (Complete A - Z Psychology by Mike Cardwell). Milgram did not have an independent variable. Unfortunately, the use of Milgram therefore resulted in many candidates scoring no marks.

# SECTION C, IN RELATION TO THE BEHAVIOURIST PERSPECTIVE

## QUESTION 17(d)

### Question 17(d)

Discuss strengths and weaknesses of laboratory experiments. Support your discussion with evidence from any studies that can be viewed from the behaviourist perspective.

[12]

## SAMPLE ANSWERS

(a) One strength of laboratory experiments is that they often have in place lots of controls to prevent the influence of extraneous variables. This allows researchers to establish cause and effect. This is seen in Bandura's study as children were given pre-existing aggression scores, then a matched pairs design was used to make sure no one group had more aggressive children than another so Bandura was increasingly sure that the results in relation to the children's aggressive acts were due to the effect of the IVs -aggressive/non-aggressive/no model.

However a weakness of laboratory experiments is that due to the strict contrived nature, results are often lacking in ecological validity. In Bandura's study children would not normally be playing alone whilst an aggressive adult bashed a bobo doll, so their behaviour cannot be generalised to other situations. Piliavin's study shows how ecological validity can be gained by using a natural instead of a laboratory setting.

Another strength of laboratory experiments is that they have highly standardised procedures that means the study can be replicated to test the external validity. This is seen in Bandura's study as each stage, witnessing a model, aggression arousal, observed play are distinct and clear in terms of what was done. Experiments are also standardised which means we can be increasingly sure that participants are not being influenced by experimenter bias, increasing the validity of the results further.

One other weakness of laboratory experiments is that they may require lots of resources and money in order to make a laboratory available. This means other aspects of the study might have to be compromised e.g. sample size in order to allow for the expenditure of the laboratory. In Piliavin's study it is clear that experimenters could freely use a public place, the train, and so were able to use all available funds to perfect the procedure and make it believable. The lack of money needed is attractive when requesting funding and so researchers are more likely to successfully get this allowing the study to take place. Because of the laboratory money might have been low, hence why only 6 participants ended up in each experimental condition, reducing population validity.

**Mark** – Third band, 7-9 marks

(b) A strength of laboratory experiments is that they are highly controlled therefore the results are reliable. For example, Bandura conducted lab experiment and controlled extraneous variables by matching participants on their natural aggression levels using the nursery teacher and an experimenter. This increases the reliability of the results and means that they can be compared against other similar experiments or to show development through age.

A weakness of lab experiments is that they are low in ecological validity because they are highly controlled and not in a natural setting. For example, Bandura's study was carried out in a lab setting as opposed to a normal nursery setting. For example, all the toys were laid out in a specific place for each child and they all had the same toys to play with. This is unlikely to occur in real life. Participants may act differently because they are not in a natural environment.

This reduces the validity of the results and reduces the usefulness of results as they cannot be used for practical applications.

A further strength of lab experiments is that they allow for the IV to be easily manipulated so cause and effect can be inferred. For example, Bandura manipulated the sex of the child, sex of model and behaviour of the model. This allowed him to conclude that it was the behaviour of the model that caused the child to act aggressively. This provided useful results that allow for practical applications such as making adults aware that their behaviour can affect their child's.

**Mark** – Third band, 7-9 marks

(c). A strength of a lab experiment is that it is high in controls. For example Samuel and Bryant and Bandura who did lab experiments are easy to replicate which makes it more reliable. They are more able to isolate the variables and conclude that it is for example the model that is affecting the child's aggression and nothing else. However it is not a natural situation and that could possibly affect how the children did on the tasks. They may have reacted differently in their normal environment. It is low in ecological validity, unlike Freud's study which was a quasi experiment. That was high in ecological validity. A lab experiment can change people's normal patterns and the results may be hard to generalise.

**Mark** – Bottom band, 1-3

## COMMENTARY

(a) Strength 1: identified (control), justified/explained (cause and effect can be inferred) and supported by appropriate evidence from Bandura. Weakness 1: (lack of ecological validity due to contrived situation), just explained/justified - through inference - (not a natural situation so behaviour cannot be generalised to other situations) supported by appropriate evidence from Bandura. Strength 2: identified (standardised procedures), justified/explained (replicability) supported by rather vague evidence from Bandura. Weakness 2: identified (often expensive as laboratories and resources may have to be hired), justified (sample size may have to be compromised) but the evidence – only because examiners know the studies – from Bandura is extremely weak and the issue has lost direction by diverting off to a strength of Piliavin's field experiment.

Therefore, although there is a range of both appropriate strengths and weaknesses, the response is not well-balanced. Understanding is evident but the answer is, at times muddled and lacks clarity. It was placed in the third band.

(b) Strength 1: identified (high control), justified (results are reliable) and supported by appropriate evidence from Bandura. Weakness 1: identified (lack of ecological validity), (not natural/normal behaviour will be shown) and supported by appropriate evidence from Bandura. Strength 2: (allow for the manipulation of IV so cause and effect can be inferred), justified (useful applications in real life) and supported by appropriate evidence from Bandura.

Weakness 2: missing!

Two appropriate strengths and one appropriate weakness have been identified, justified/explained and supported by appropriate evidence. The response was therefore placed in the third band.

(c) Strength 1: identified (control) and justified (replicability and reliability) with weak evidence from Bandura (S&B not acceptable). The answer then becomes completely muddled and mostly irrelevant so was placed in the bottom band.

### How the answer could be improved

One strength of laboratory experiments is that the procedure is standardised. This means the experiment can be replicated to check for reliability. For example the study by Bandura, when in the first room, the aggressive model acted in an aggressive manner by abusing the bobo doll in a standardised way both verbally and physically and in the last (experimental) room the toys, including a 3ft bobo doll, a mallet, a tea set and plastic farm animals were always arranged in a fixed order. This procedural standardisation allowed for easy replication of the experiment and showed findings to be reliable / consistent.

A second strength of laboratory experiments is they allow for the manipulation of an independent variable to see the effect on a dependent variable. This means one can infer cause and effect. For example Reicher and Haslam manipulated the variable of permeability by making it possible for one prisoner to be promoted to the role of guard to see the effect this had on the prisoners' behaviour. This led to the prisoners not identifying as a group and working independently to demonstrate the qualities required to become a guard. This allowed the researchers to infer that as long as individuals have the opportunity for social advancement they will act independently and not develop a group identity or social cohesion.

One weakness of laboratory experiments is that they lack ecological validity. This means they may not reflect real life situations or how people may behave in natural environments. For example Reicher and Haslam's experiment was not only conducted in a specially constructed mock prison at Elstree Film Studios in London, therefore a laboratory, but also the participants who were randomly assigned to the roles of prisoners and guards were not genuine prisoners or guards so this did not reflect real life. Real prisoners and guards, in a real prison may behave differently.

Another weakness of laboratory experiments is that because participants are usually aware that they are taking part in an experiment they may respond with demand characteristics. This is where participants work out the aim of the study and so behave in ways they think the researcher wants them to, rather than behaving naturally. This therefore makes the validity of the findings questionable. For example the participants in Reicher and Haslam's prison study knew they were not only taking part in an experiment but that they were also being filmed for television. The prisoners in particular may therefore have behaved in extreme stereotypical ways as that's how they thought the researchers would have expected them to behave e.g. prisoner JE throwing his lunch plate to the ground demanding better food and the prisoners stealing the keys from the guards.

### **General performance on the question**

awarded. As in previous sessions some of the strengths/weaknesses were merely identified, not identified and justified/explained e.g. a weakness of laboratory experiments is that they lack ecological validity – so? Marks were generally lost in either of two ways:

(i) Using inappropriate supporting evidence.

(ii) Producing a response that related specifically to the behaviourist perspective and not laboratory experiments i.e. producing an answer 'I prepared earlier'.

# SECTION C, IN RELATION TO THE COGNITIVE APPROACH

## QUESTION 18(a)

### Question 18(a)

Briefly outline the cognitive approach.

[2]

## SAMPLE ANSWERS

(a) The cognitive approach assumes that internal mental processes such as language and thinking are important influences on our behaviour.

**Mark – 2 marks**

(b) The cognitive approach focuses on mental processes and how the brain inputs, processes and outputs information like a computer. The processes focused on include memory and language.

**Mark – 1 mark**

(c) The cognitive approach sees us as computers.

**Mark – 0 marks**

## COMMENTARY

(a) This is an appropriate outline that is linked to both aspects of the cognitive approach and behaviour. It is worthy of being awarded the full 2 marks.

(b) There is no link to behaviour and so this is only a partial response and only 1 mark could be awarded.

(c) This is far too vague to be creditworthy. 0 marks were awarded.

### How the answer could be improved

- The cognitive approach assumes our internal mental processes such as memory, problem solving, reasoning, influence our behaviour.
- The cognitive approach sees the human mind as working like a computer – it inputs, processes and responds to and behaviour is influenced by the way information is received and processed.

### General performance on the question

This was generally well answered. Some candidates used the computer analogy but failed to link this behaviour thus losing valuable marks.

# SECTION C, IN RELATION TO THE COGNITIVE APPROACH

## QUESTION 18(b)

### Question 18(b)

Describe how the cognitive approach could explain why individuals with autism have difficulty understanding other people's mental states.

[4]

## SAMPLE ANSWERS

(a) The cognitive approach could explain why individuals with autism have difficulty understanding other people's mental states due to them having an impaired theory of mind, meaning they find it hard to 'read people's minds' or understand their feelings – they cannot put themselves in other people's shoes. Baron-Cohen did a piece of research into autistic people in which they scored significantly less in an Eyes Task, a test of theory of mind than 'normal' people or people with Tourette syndrome. The autistic people were unable to read the emotions shown in people's eyes and scored an average of 16.3/25 on the Eyes Task compared to 20.3/25 scored by 'normal' people and 20.4/25 by those with Tourettes. This shows that the cognitive approach can explain how people with autism have difficulty understanding other people's mental states as they have the cognitive deficit of a poor theory of mind.

**Mark** – Top band, 3-4 marks

(b) The cognitive approach believes that our behaviour is influenced by internal mental processes. This can be used to explain why people with autism have difficulties understanding other people's mental states as they have a lack of theory of mind. This is explained by this approach as people with autism have issues with how they process and perceive information. For example in Baron-Cohen's study of autism, the autistic and AS group was significantly impaired in the Eyes Task at understanding people's emotions by 'mind reading' compared to the 'normal' and Tourettes groups, so their lack of theory of mind means they cannot put themselves in some else's shoes and understand how they are feeling.

**Mark** – Top band, 3-4 marks

(c) Baron-Cohen carried out a study to see why autistic people have difficulties understanding other people's mental states. He found that in the 'Eyes Task' the autism group got a mean score of 16.3 whereas the 'normal' group scored 20.4. This showed that autism affects the TOM of a person. Theory of mind is when someone can put themselves in another person's situation and understand that people have different thoughts and opinions to others.

**Mark** – 2 marks

## COMMENTARY

(a) This description is accurate and shows good understanding. There are clear links to the approach and the elaboration in relation to the Baron-Cohen's study is appropriate. The answer was placed in the second / top band.

(b) Although the literacy could have been better, this description shows a clear link between the cognitive approach and the focus of the question. Appropriate evidence from Baron-Cohen has been cited and understanding is evident. The answer was placed in the second / top band.

(c) Although some understanding is evident, this answer is rather muddled and, in places inaccurate e.g. the 'normal' participants scored an average of 20.3/25 in the Eyes task, not 20.4. The candidate has explained what TOM is but has failed to explain its relevance to this particular situation. The answer was awarded 2 marks and placed in the bottom band.

**How the answer could be improved**

An assumption of the cognitive approach is that internal mental processes influence the way people behave. It has been found that many autistic people have impairments in the development of a theory of mind – a cognitive process that enables people to attribute mental states (knowledge, beliefs, desires and feelings) to others. Lacking a theory of mind may be the core deficit that underlies the social, communicative and imaginative dysfunctions of autistic people. Theory of mind is one aspect of autistic mind-blindness and deficits in the normal process of empathising, which also involves appropriate emotional reactions (e.g. sympathy) to other people. Baron-Cohen found that autistic adults had tremendous difficulty in recognising other people's emotional states. Using the Eyes Task, designed to be a 'pure' theory of mind test, he found that autistic participants had more difficulties than either normal adults or adults with Tourette's syndrome in recognising mental states from photographs of eyes – their mean score was 16.3/25 compared to 20.3/25 for normal adults and 20.4/25 for those with Tourette's. This indicated that the autistic participants lacked a theory of mind. This deficit in mental processing may therefore explain why such people have difficulty empathising with others / understanding other people's mental states.

**General performance on the question**

Overall, this question was better answered than in previous sessions. Many candidates were able to describe appropriate evidence (Baron-Cohen)/an appropriate example to show why individuals with autism have difficulty understanding other peoples' mental states. Links to the approach were better than in previous sessions though many candidates didn't adequately relate theory of mind to the cognitive approach.

# SECTION C, IN RELATION TO THE COGNITIVE APPROACH

## QUESTION 18(c)

### Question 18(c)

Describe one similarity and one difference between any experiments that can be viewed from the cognitive approach. [6]

## SAMPLE ANSWERS

(a) One similarity between the experiments by Loftus and Palmer and Baron-Cohen is that there were both highly controlled. Loftus and Palmer controlled the number and content of the videos shown to participants so they all saw the same thing and Baron-Cohen controlled the pictures shown to participants so they all saw the same 25 black and white photos of people's eyes. This level of control results in high reliability.

One difference between the experiments of Loftus and Palmer and Baron-Cohen is the sample size. Loftus and Palmer had 45 participants in experiment 1 and 150 in experiment 2 (total 195) whereas Baron-Cohen had 50 'normal' participants, 16 autistic/AS participants and 10 participants with Tourette syndrome (total 76).

**Mark – 6 marks**

(b) A similarity between Loftus and Palmer's study and Baron-Cohen's study is that they both collect quantitative data. Loftus and Palmer collected quantitative data through the estimated speed of the cars from each group in experiment 1 as well as estimated speeds in experiment 2 and the amount of people who claimed to have seen broken glass in the car crash in experiment 2. Baron-Cohen gathered quantitative data through collecting how many times the participants (autistics/'normal'/Tourettes) answered correctly / how many errors they made. This was collected from the eyes task, gender recognition task, the basic emotions task and Happé's strange stories.

A difference is that Loftus and Palmer's experiment was a laboratory experiment whereas Baron-Cohen's was a quasi experiment. Loftus and Palmer's study was a laboratory experiment because it contained an IV that was not naturally occurring and a DV. However in Baron-Cohen's study the IV was naturally occurring. The IV was autism/Tourettes/'normal', all of which could not be manipulated by the researcher.

**Mark – 5 marks**

(c) One similarity between Baron-Cohen and Loftus and Palmer is that participants were humans. In Baron-Cohen, 16 adults with high-functioning autism or Asperger's syndrome, 10 adults with Tourettes and 50 'normal' adults were used. In Loftus and Palmer 45 students were used in experiment 1 and 150 students were used in experiment 2.

One difference between Baron-Cohen and Savage-Rumbaugh is that they investigated different things. Baron-Cohen investigated whether autistic people lack a theory of mind whilst Savage-Rumbaugh investigated the spontaneous language capabilities in pygmy chimpanzees.

**Mark – 3 marks**

(d) A similarity between Baron-Cohen and Savage-Rumbaugh is that they are both quasi experiments. The IV in Baron-Cohen is whether the participants has Tourettes, autisms/Asperger's or were 'normal', and the IV in Savage-Rumbaugh was whether the chimps were common chimps or pygmy chimps. Neither of these studies could manipulate their IV.

A difference between Loftus and Palmer and Savage-Rumbaugh was the size of the samples. In their first study Loftus and Palmer had a sample of 45 participants and in the second experiment they had a sample of 150 participants. Savage-Rumbaugh only had a sample of 4 in total: 2 pygmy chimps and 2 common chimps.

**Mark – 0 marks**

## COMMENTARY

(a) Similarity: an appropriate similarity has been identified (both were highly controlled). This has then been supported by appropriate evidence from both Loftus and Palmer and Baron-Cohen.

Difference: an appropriate difference has been identified (different sample sizes). This has then been supported by appropriate evidence from both Loftus and Palmer and Baron-Cohen.

The response was awarded  $3+3=6$  marks.

(b) Similarity: an appropriate similarity has been identified (both collected quantitative data). This has then been supported by appropriate evidence from both Loftus and Palmer and Baron-Cohen.

Difference: an appropriate difference has been identified (Loftus and Palmer manipulated their IV whereas Baron-Cohen had a naturally occurring IV). This has then been supported by appropriate evidence from Baron-Cohen but inadequate evidence from Loftus and Palmer i.e. the candidate has failed to identify the manipulated IV. The response was therefore awarded  $3+2=5$  marks.

(c) Similarity: an appropriate similarity has been identified (both used human participants). This has been supported by appropriate evidence from both Baron-Cohen and Loftus and Palmer. 3 marks

Difference: this is not creditworthy as the candidate has used an inappropriate study – Savage-Rumbaugh did not conduct an experiment. 0 marks. The response therefore gained  $3+0=3$  marks.

(d) No credit can be given to either the similarity or the difference as Savage-Rumbaugh's study was not a laboratory experiment. 0 marks were awarded.

### How the answer could be improved

One similarity between two experiments that can be viewed from the cognitive approach is that both of the experiments conducted by Loftus and Palmer and the experiment conducted by Samuel and Bryant collected quantitative data. Loftus and Palmer collected quantitative data through the estimated speed of the cars from each group in experiment 1 as well as estimated speeds in experiment 2 and the amount of people who claimed to have seen broken glass in the car crash in experiment 2. Samuel and Bryant collected quantitative data in relation to whether or not children aged 5,6,7 and 8 could / could not conserve number, mass and volume in the three conditions of standard, two questions, one question and fixed array.

One difference between two experiments that can be viewed from the cognitive approach is the sample used. Loftus and Palmer used 45 students in experiment 1 and 150 students in experiment 2 whereas Samuel and Bryant used 252 boys and girls between the ages of 5 and  $8\frac{1}{2}$  years.

### General performance on the question

Many candidates failed to respond to the question properly and their answers related to the cognitive approach, rather than experiments that can be viewed from the cognitive approach. Furthermore, many candidates compared inappropriate studies because they did not seem to be fully aware of the features of an experiment. Savage-Rumbaugh is not an experiment because the pygmy and common chimpanzees were not reared or tested under the same conditions. The study is generally considered a longitudinal case study. Unfortunately, the use of Savage-Rumbaugh therefore resulted in many candidates scoring no marks.

# SECTION C, IN RELATION TO THE COGNITIVE APPROACH

## QUESTION 18(d)

### Question 18(d)

Discuss strengths and weaknesses of laboratory experiments. Support your discussion with evidence from any studies that can take the cognitive approach. [12]

## SAMPLE ANSWERS

(a) A strength of laboratory experiments is that they have high control over confounding variables. In Loftus and Palmer's study, the 7 video clips of car accidents were standardised and the questions in the questionnaire given (except for the critical question) were all standardised and the same. By standardising the clips seen and the questions it means the speed estimates given are only influenced by the verb in the critical question rather different content of video clip and different questions given. It's more likely only the independent variable is affecting the dependent variable so internal validity is high and it's easier to make cause and effect statements.

Another strength is that laboratory experiments are easy to replicate. In Baron-Cohen's study, the photographs in the eyes task were standardised for all participants, so the same material can be used for every participant as the procedure is identical. As the procedure is identical for all groups – autistics/AS adults, Tourette syndrome adults and 'normal' adults, it's easy for the researcher to replicate for all participants. As it's easy to replicate the results produced are more reliable.

A weakness of laboratory experiments is that there is low ecological validity. For example, to test theory of mind in Baron-Cohen's study, it required that the participants identified the emotion shown by a pair of black and white eyes that were projected for 3 seconds. This task is unrealistic because one does not look at a person's eyes in black and white. So the score of their theory of mind may not be representative of what their theory of mind is like in a natural situation. By using a laboratory experiment results are hard to generalise to real life because laboratory experiments are not representative of real life situations so results can't be used to explain behaviour in normal situations.

Another weakness of laboratory experiments is that they can put participants under stress. In Baron-Cohen's study participants had only 3 seconds to identify the emotion shown in the eyes task. Because the participant had a time limit to see the pair of eyes this could have caused a lot of pressure on the participants leading to stress. Because laboratory experiments are conducted in controlled conditions, pressure and stress can be put on participants. They are not in their natural environment so it's unethical.

**Mark – 12/12 marks**

(b) A strength of lab experiments is that they are highly controlled which increases their reliability and decreases the chances of extraneous variables influencing results. An example of a study which is high in control is the Baron-Cohen study. Each participant was tested individually and was shown an image which was 15 x 10 cm. (25 in total, all black and white, including men and women) for 3 seconds. Hence, as all participants went through the same standardised and controlled procedure they are less likely to react differently due to extraneous variables so results would be more reliable and consistent.

However a weakness of the high control in lab experiments is that they aren't ecologically valid. For example Loftus and Palmer in experiment 1, showed all their participants 7 film clips of car crashes. Seeing crashes in real life would be different to seeing them in clips, so participants may react in different ways than if they saw a real car accident – hence the study does not reflect real life.

Another weakness of laboratory experiments is that they have small samples. An example of a study from the cognitive approach with a small sample is Baron-Cohen: there were only 10 people in the Tourettes group and 16 in the autistic/Asperger's group (12 of whom had Asperger syndrome and 4 of whom had autism). Small samples tend to lack representativeness of the rest of the population or of a larger group. Hence it decreases the generalisability of the results.

(b) cont.

Another strength of laboratory experiments is that they tend to collect quantitative data. For example Loftus and Palmer collected data on the average speed of the cars estimated by all participants in each condition (e.g. the people in the 'smashed' condition reported higher speeds than those in the 'contacted' condition). Quantitative data is easily comparable and easy to analyse and more likely to be reliable than qualitative data, making the results more valid.

**Mark** – Third band, 7-9 marks

(c) One strength of a laboratory experiment is that it is often highly controlled, meaning they have a high level of reliability and validity due to the extraneous variables being easier to control. For example, in Loftus and Palmer's study on eye witness testimony, the participants all watched the exact same video clips, for the exact same amount of time. This means that nothing is different: cars travelled at the same speed, in the same setting etc. This is a strength because it means data is likely to be more reliable and accurate.

However one weakness of a lab experiment is that it can be very low in ecological validity. For example, in the same study by Loftus and Palmer, watching a real car crash may be much more traumatic and perhaps increase memory quality and mean that leading questions don't have much of an effect on recall. This is a weakness because it means that results are not totally applicable to real-life situations and to actual human behaviour.

**Mark** – Second band, 4-6 marks

## COMMENTARY

(a) Strength 1: identified (high control), justified/explained (cause and effect can be inferred) supported by appropriate evidence from L&P; Strength 2: identified (replicability), justified (reliability) supported by appropriate evidence from B-C; Weakness 1: identified (lack of ecological validity), justified (doesn't represent real life situations) supported by appropriate evidence from B-C; Weakness 2: identified (stress) and justified (unethical) supported by appropriate evidence from B-C. Full marks awarded. Two appropriate strengths and two appropriate weaknesses have been identified, explained/justified and supported with appropriate evidence. The response was therefore placed in the top band.

(b) Strength 1: identified (high control), justified (reliability), supported by appropriate evidence from Baron-Cohen; Weakness 1: identified (low ecological validity), justified (does not reflect real life), supported by appropriate evidence from Loftus and Palmer; Weakness 2: not creditworthy as lab experiments frequently have large samples and even Baron-Cohen used 50 'normal' participants in his study; Strength 2: identified (quantitative data), justified (easy to compare/analyse, more reliable), supported by appropriate evidence from Loftus and Palmer. There are therefore two creditworthy strengths and one creditworthy weakness which allowed the answer to be placed in the third band.

(c) Strength 1: identified (high control), explained (reliability and validity), supported by appropriate evidence from Loftus and Palmer. Weakness 1: identified (low in ecological validity), justified/explained (behaviour does not reflect what would happen in real life) supported by appropriate evidence from Loftus and Palmer.

Unfortunately it appears this candidate ran out of time. One strength and one weakness have been identified, justified and supported by appropriate evidence. As there is not a range of strengths and weaknesses, the response could therefore not score more than 6 marks, so it was placed in the second band.

### How the answer could be improved

One strength of laboratory experiments is that the procedure is standardised. This means the experiment can be replicated to check for reliability. For example in the study by Loftus and Palmer's first experiment all participants saw exactly the same seven film clips of car crashes and were given the same questionnaire to complete with the exception of the question containing the critical verb relating to speed. Likewise, in the second experiment, all participants watched the same film clip of the car accident, all completed the same questionnaire with the exception of the question with/without the critical verb relating to speed, all returned one week later, and all completed the same second questionnaire that contained the critical question relating to broken glass. This procedural standardisation allowed for easy replication of the experiment and showed findings to be reliable / consistent.

A second strength of laboratory experiments is they allow for the manipulation of an independent variable to see the effect on a dependent variable. This means one can infer cause and effect. For example in their first experiment Loftus and Palmer manipulated the verb in the critical question, 'How fast were the cars going when they smashed / collided / hit / contacted / bumped into each other?' to see the effect of a verb on speed estimates. This allowed the researchers to infer that the wording of a question can influence the response given and that leading questions can affect the accuracy of eyewitness testimony.

One weakness of laboratory experiments is that they lack ecological validity. This means they may not reflect real life situations or how people may behave in natural environments. For example in Baron-Cohen's study participants, when doing the Eyes task, were shown an image which was 15 x 10 cm. (25 sets of eyes in total, all black and white, including men and women) for 3 seconds. After this they had to decide, from a choice of two alternatives, what emotion was being displayed in the eyes. This did not reflect real life as normally one would judge emotion from looking at the eyes (or even the whole face / body) of a live person.

Another weakness of laboratory experiments is that because participants are usually aware that they are taking part in an experiment they may respond with demand characteristics. This is where participants work out the aim of the study and so behave in ways they think the researcher wants them to, rather than behaving naturally. This therefore makes the validity of the findings questionable. For example the participants in Loftus and Palmer's experiments may have realised that they were expected to give a speed appropriate to the verb given in the critical question and so gave speed estimates they thought Loftus and Palmer would expect rather than the speed they really thought the cars were travelling at. Likewise, in the second experiment participants who had been asked how fast the cars were going when they smashed into each other may have realised that they would be expected to say they had seen broken glass and so responded accordingly when actually they may have thought they had no broken glass.

### **General performance on the question**

There were many good answers with the full range of marks being awarded. As in previous sessions some of the strengths/weaknesses were merely identified and not justified or explained e.g. a weakness of laboratory experiments is that they lack ecological validity, with nothing further. Marks were generally lost in either of two ways:

- (i) Using inappropriate supporting evidence.
- (ii) Producing a response that related specifically to the cognitive approach and not laboratory experiments i.e. producing an answer 'I prepared earlier'.

## Contact us

Keep up to date with the latest news by registering to receive e-alerts at [www.ocr.org.uk/updates](http://www.ocr.org.uk/updates)

Telephone 01223 553998

Facsimile 01223 552627

Email [general.qualifications@ocr.org.uk](mailto:general.qualifications@ocr.org.uk)

