Cambridge International AS & A Level

Cambridge International Examinations
Cambridge International Advanced Subsidiary and Advanced Level

PSYCHOLOGY **9698/12**

Paper 1 The Core Studies 1 **May/June 2016**

MARK SCHEME

Maximum Mark: 80

Published

This mark scheme is published as an aid to teachers and candidates, to indicate the requirements of the examination. It shows the basis on which Examiners were instructed to award marks. It does not indicate the details of the discussions that took place at an Examiners' meeting before marking began, which would have considered the acceptability of alternative answers.

Mark schemes should be read in conjunction with the question paper and the Principal Examiner Report for Teachers.

Cambridge will not enter into discussions about these mark schemes.

Cambridge is publishing the mark schemes for the May/June 2016 series for most Cambridge IGCSE®, Cambridge International A and AS Level components and some Cambridge O Level components.

1 In the study by Loftus and Pickrell, one way that false memories were tested was by the number of words the participants used to describe the memories.

(a) Describe the results obtained for the number of words. [2]

Mean number of words used:
True: 138
False: 49.9

6/7 participants used more words to describe the true stories than the false ones.
1/7 participants used very few words to describe either true (mean 20) or false (mean 21) stories.

1 mark partial ('true more words (than false)' and/or one figure correct)
2 marks full (both figures approximately correct)

(b) Suggest <u>one</u> advantage of using a word count as a measure of false memory. [2]

It is a quantitative measure, so is objective/cannot be affected by the opinion/interpretation of the researcher (2 marks)

It is a more informative than simply whether they remember or not, as it allows for a range of possible answers/it tells you the *extent to which* they remembered (it is an ordinal scale rather than simply a nominal one).

1 mark partial (simple explanation of an advantage)
2 marks full (elaborated explanation)

2 Held and Hein conducted an investigation into visual development using animals in a carousel apparatus.

(a) Describe the sample of animals used. [2]

Kittens/cats
Neonates/newborn
10 pairs
Each pair from a different litter
(Reared in darkness from birth then) aged 8–12 weeks started the experiment
(Group Y started at 2 weeks of age)

1 mark per idea, e.g. '10 pairs of kittens/cats' = 2 marks.

NB Accept 8 pairs (i.e. group X only)

(b) Suggest why this species was chosen rather than any other non-human animal. [2]

Practical:
Open their eyes soon after birth; so we can distinguish nature and nurture;
They are able to move when very young
Small enough to fit into such apparatus
Strong enough to move such apparatus
They have observable behaviours; that indicate whether they can perceive depth or not;
Similar to humans; because they are mammals;

Ethical:
Possible to deprive them (which might not be possible for higher animals such as great apes)

1 mark partial (simple suggestion for choice)
2 marks full (suggestion with clear explanation)

NB Accept any reasonable suggestion.

3 From the study by Baron-Cohen et al. (eyes test):

(a) Name and outline the experimental design used. [2]

Independent groups/between participants (subjects); any one participant performs in only one of the levels of the independent variable/different groups are used in different levels of the independent variable.

1 mark = name of design (i.e. independent groups/between participants (subjects))
1 mark = outline of design (this may be contextualised, but does not have to be)
1 mark = weak naming (e.g. 'independent') plus weak outline

NB Do not accept just 'independent' or 'between' (because 'independent' could just mean 'independent variable', which is incorrect)
NB 'All participants do all tasks/tests' is **not** correct as this is repeated measures.

(b) Explain why the design that you named in (a) was used, instead of an alternative experimental design. [2]

Essential – participants were either autistic or not = 2 marks
Good because able to match for IQ; across groups 1 and 4;
Repeated measures could not be used; as participants could not be both control and autistic;
Matched pairs hard to use; as matching on relevant variables such as IQ would be difficult/ hard to decide what variables it would be important to match on;

1 mark partial (naming alternative: repeated measures/matched pairs or simple comment)
2 marks full (detailed comment e.g. alternative plus why it was not chosen – any reasonable explanation earns credit)

4 From the prison simulation by Haney, Banks and Zimbardo:

(a) Name and describe the sampling method used in this study. [2]

Volunteer/self-selected sampling.
Participants respond to an advert or request/e.g. through newspaper advertisement in this study.

Naming type of sampling = 1 mark
Explanation with or without contextualisation = 1 mark

NB newspaper/advertisement alone = 0 marks

(b) Give <u>one</u> advantage of this sampling method in this study. [2]

Most likely:
May be representative of wider population, if a wide sample is available through the advert; because mainly male prisoners/because lots of people read newspapers (any link to study); can obtain sample with specific characteristics;

1 mark partial (statement of advantage, however brief)
2 marks full (application to HBZ, however brief).

5 From the study by Piliavin et al. (subway Samaritans):

(a) Describe <u>one</u> difference between the results from the drunk and ill victims. [2]

Most likely
The ill (cane) victim was helped more (than the drunk victim) (1 mark) but only when the victim was black (2 mark).
Drunk victims were more likely to be helped by their own race (2 mark).
More people left the critical area with drunk victims than ill (1 mark).
More people made comments with drunk victims than ill (1 mark).

100% for cane white and black, and for drunk white but 73% for drunk Black (2 marks)

1 mark partial (difference stated)
2 marks full (some detail/data)

(b) Explain why there was a difference between the results from the drunk and ill victims. **[2]**

The explanation will depend on the evidence given.

Most likely
Sympathy was greater for the ill victim
The drunk was perceived as dangerous; so less helping/more leaving
Blame/fear/disgust could prevent helping for the drunk victim
Empathy/trust are greater for same race so more likely to help (when risks are high)/costs-benefits
Prejudice is greater for different race in the drunk condition
Discomfort in the drunk condition made people talk more; they felt the need to explain why they weren't helping

1 mark partial (simple explanation)
2 marks full (some detail)

6 The study by Tajfel (intergroup categorisation) was a laboratory experiment.

(a) Describe <u>two</u> features shared by all laboratory experiments. **[2]**

- IV and DV
- Control (of variables)
- IV manipulated
- DV measured
- artificial environment
- comparison between groups
- looking for differences
- investigates causal relationships

1 mark partial (neither of first two points)
2 marks full (has one of first two points plus other relevant information)

(b) Explain why study 1 from Tajfel was an experiment. **[2]**

Controlled environment: where the boys' groups were allocated; rather than real groups which might have real differences;
IV: in-group/out-group; and *DV*: allocation of points.

1 mark partial (brief explanation related to the study)
2 marks full (detailed explanation related to the study)

NB No marks for just saying has IV and DV, answer must be related to study

7 From the study by Bandura et al. (aggression):

(a) Describe what is meant by a repeated measures design. [2]

An experimental design in which any one participant/group performs in each of the conditions/levels of the independent variable (i.e. would have seen both male and female model/ would have seen no model, aggressive model or non-aggressive model).

1 mark partial = a correct but unclear description
2 marks full = a correct and clear description (this may be contextualised, but does not have to be)

NB 'All participants do all tasks/tests' is **not** correct as this may also be true of independent groups (it is the *conditions* or *levels of the IV* which matter).

(b) Explain a practical problem if a repeated measures design had been used in this study. [2]

If a participant had seen both a male and a female model/ no model, aggressive model and non-aggressive model they would have had more exposure to the model so might have responded differently because of this, rather than because of the model's behaviour itself (which would have reduced validity/created order effects).

1 mark partial = simple or muddled explanation which is not contextualised
2 marks full = a clear explanation which is contextualised

8 From the study by Freud (little Hans):

(a) Describe the research method used. [2]

Case study;
Studies one individual (or one instance, e.g. a family);
Studied in depth/detail/descriptive;
May use many techniques (e.g. observations, questionnaires, interviews, tests) looking for interactions, complex relationships

1 mark partial = simple or muddled description
2 marks full = detailed description (does not require contextualisation).

(b) Explain why this was the best choice of research method for this study. [2]

Allowed for study of Hans in detail;
To look for links between his behaviour/dreams/fantasies and the Oedipus complex;
This would not have been possible in a study investigating only one or two variables (such as an experiment);

1 mark partial = simple or muddled explanation
2 marks full = clear explanation (with reference to little Hans).

9 From the study by Langlois et al. (infant facial preference):

(a) Identify any <u>two</u> variables that were manipulated. [2]

Preference for attractive faces in:
Gender: (White) adult male; versus (white) adult female faces
Race: Black adult female faces (compared to white adult female faces)
Age: infant faces (compared to adult faces)

Name of IV/two conditions of one IV = 1 mark × 2

The only IVs were:
Gender/males **and** females
Age/babies **and** adults
Race/Black **and** White
Any 2 = 2 marks

(b) Identify the <u>two</u> ways in which Langlois et al. concluded that infant facial preferences could be acquired. [2]

innate/inborn/biological/genetic/inherited/evolutionary/through 'nature';
learned/environmental/cultural/through 'nurture';

1 mark per way named × 2
1 mark 'nature', 1 mark 'nurture'

NB Do **not** accept 'natural' for innate

10 Describe <u>two</u> of Nelson's conclusions from the study on children's morals. [4]

Main conclusions:
morality develops (7 years olds better than 3 year olds)
Children develop morality quite early (3 year olds could make moral judgments)

Specific conclusions:
Making a moral judgment requires understanding of the evaluative concepts i.e. good and bad to be applied;
A comprehension of the motives or goals involved e.g. sharing, helping, hurting;
A child must also be able to recognise and interpret the interrelationship between actions, motives, goals, and outcomes in order to make evaluative judgments;
Children as young as 3 years of age can and do use motive information for making moral judgments; when this information is explicit and salient;
When motive and outcome have opposite valences, children tend to recall the story so as to make them congruent; so may make moral misjudgments.

1 mark partial (brief description)
2 marks full (some detail) **per conclusion × 2**

11 The study by Schachter and Singer investigated emotion. Describe why the study was conducted. [4]

Most likely
To test the two-factor theory (of emotions);
which says that emotions are governed by physiological and psychological factors;
physiological factors are physical/biological;
psychological factors (in this study) were cognitive;
the physiological factor was adrenalin (epinephrine);
which is a hormone that increases arousal;
the psychological factor was created by the stooge;
either anger or happiness;
and was manipulated by information;
informed, misinformed, ignorant;
cognitive labelling of arousal to understand;

1 mark per detail as above × 4

12 From the study by Demattè et al. (smells and facial attractiveness):

(a) Describe <u>one</u> control used in the study. [2]

same flow rate; controlled by olfactometer;
same presentation duration; 500ms;
odour dilution; controlled so that all had the same perceived intensity;
participants' health; asked about colds etc.;
faces counterbalanced across conditions; i.e. different faces with different smells, but each group of faces seen with each smell by some participants;
clean air between trials; through olfactometer;
chair/chin rest; participant same distance from screen;

1 mark partial (brief/identified)
2 marks full (detail)

(b) Explain why failing to control this variable would have been important in this study. [2]

(same flow rate): faster smell might be more obvious; so some participants might notice it more easily;
(odour dilution): stronger smell would be more obvious; so some participants would notice it more easily;
(participants' health): having a cold blocks your nose; so you'd notice the smells less/respond more to the image;
(faces counterbalanced across conditions): some faces might have matched /clashed with the smells; making the effect of the smell greater/less;
(clean air between trials): the clean air removes any residual smell; so the new smell wasn't masked by the last one;
(chair/chin rest): if some participants were closer to the screen they might see the faces more clearly; so be more affected by the perception/less affected by the smell;

1 mark partial (brief explanation)
2 marks full (some detail)

13 In the study by Rosenhan (sane in insane places), qualitative data were collected.

(a) Describe <u>two</u> pieces of qualitative data from the study. [2]

Staff behaviour towards patients:
- kept away from them (stayed in staff room/spent little time with them)
- when asked questions: answered briefly/with head averted/they often didn't respond at all
- were rude/inappropriate to patients

Staff interpretation of pseudo-patient behaviour:
- queuing early for food: oral-acquisitive behaviour (symptom of disorder/rather than boredom)
- note-taking: engaging in writing behaviour (symptom of forgetting/compulsion of schizophrenia rather than recording events)
- walking corridors: nervous behaviour (rather than boredom)

Patient interpretation of pseudo-patient behaviour:
- thought they were journalists
- thought they were professors
- thought they were checking up on the hospital
- often insisted the pseudo-patient was sane (even when the staff did not)

1 mark per piece of qualitative data identified **× 2**

(b) Suggest <u>one</u> advantage of collecting qualitative data in this study. [2]

able to record detail; e.g. subtleties of behaviour of staff towards patients/interpretation of (ab)normal behaviour;
could record a range of different variables; staff responses, real patients' responses;

1 mark partial (advantage not related to this study)
2 marks full (advantage related to this study, however briefly)

14 In the study by Thigpen and Cleckley (multiple personality disorder), Eve reported having blackouts.

(a) Identify <u>two</u> <u>other</u> pieces of information about the participant that were known to Thigpen and Cleckley at the start of the study.

25 year old woman;
Married;
With a child;
She had severe/blinding headaches;
Emotional difficulties (complex marital problems and personal frustrations)

1 mark for any piece of information from above **× 2**

(b) Describe how Thigpen and Cleckley explained Eve's blackouts. [2]

Most likely:
The blackouts were due to a second personality; Eve Black/EB; who was in control when Eve White didn't remember.

1 mark partial (simple or muddled description)
2 marks full (clear description)

15 Describe the embedded figures test used in the study by Billington et al. (empathising and systemising). You may use a diagram in your answer. [4]

Figures had a **pair of diagrams** and a small shape;
12 of them;
ask to **find shape** (simple/small/black and white);
hidden in one of the big/complex/coloured shapes;
responded by pressing a key to indicate correct shape was on the left or right;
i.e. selecting one of two possible answers is **forced choice**;
score of **number correct** (and **speed**/timed)
maximum 50 seconds before moving on automatically
done **online**.

accept any '**shape within a shape' drawn**.

The ideas in bold are key.

0 marks: incorrect
1 mark: some relevant information
2 marks: some aspects of test are present but not sufficient to be clear
3 marks: clear outline of test or fair outline of test and of data collected
4 marks: clear outline of test and of data collected

NB Two marks maximum for good but unlabelled diagram (whether made up or as in the study)

16 Discuss the extent to which generalisations that can be made from <u>one</u> of the studies listed below.

Milgram (obedience)
Maguire et al. (taxi drivers)
Veale and Riley (mirror gazing) **[10]**

No marks for description of study.
Max 5 if only consider either generalisable/not generalisable.

Comment	mark
No answer or incorrect answer.	0
Anecdotal evaluation, brief detail, minimal focus. Very limited range. Evaluation may be inaccurate, incomplete or muddled.	1–3
Either points illustrating generalisations lack depth and/or breadth, or only strengths or only weaknesses of generalisations are considered. The answer is general rather than focused on study but shows some understanding.	4–5
Some strength(s)/weakness(es)/applications of generalisations are considered and argument is focused on the study although the evaluation may be imbalanced in terms of quality and/or depth. The answer shows reasonable understanding.	6–7
There is a range of strengths/weaknesses/applications of making generalisations, the discussion is focused on the study. Evaluation is detailed with good understanding and clear expression.	8–10

Examples of possible evaluation points:

Milgram
- *Can generalise* because many people obeyed
- Americans obeyed so can generalise beyond 'Germans are different'
- replicated widely since so can make generalisations
- *Cannot generalise* based on narrow sample originally, so initial generalisations potentially flawed
- based on lab study, how well does this extend to real life, so initial generalisations potentially flawed

Maguire et al.
- *Can generalise* because it was a study of biological processes and these tend to be similar between everyone
- There are many factors being considered – like real world navigation – so the results are complex (although some factors are deliberately not considered, e.g. the role of spatiotemporal contexts of landmarks or films)
- *Cannot generalise* because generalisations to real-world navigation from imagining routes may not be valid as memory for routes may depend on the experience of what is seen as we travel
- All the taxi drivers were male. As female navigation differs from males, the findings may not generalise to females.

Veale and Riley
- *Can generalise* because study used 52 real patients so findings should generalise to other real patients
- participants included males and females, so findings should generalise across genders
- *Cannot generalise* because most of data collected were from closed questions and from one moment in time. The findings may not be representative of changing situations and may not allow participants to express themselves fully.
- based mainly on self-report data and it is difficult to know how accurate this was (as participants may lie about their behaviour to appear socially acceptable), so generalisations potentially flawed.

17 Evaluate <u>one</u> of the studies listed below in terms of its strengths and weaknesses.

Mann et al. (lying)
Baron-Cohen et al. (eyes test)
Dement and Kleitman (sleep and dreaming) **[10]**

No marks for description of study.
Max 5 if only about being strengths or only about weaknesses.

Comment	mark
No answer or incorrect answer	0
Anecdotal discussion, brief detail, minimal focus. Very limited range. Discussion may be inaccurate, incomplete or muddled.	1–3
Either points illustrating strengths or weaknesses of the study are limited, or lack of depth and/or breadth. The answer is general rather than focused on study but shows some understanding.	4–5
Both strength(s) and weakness(es) are considered and are clearly focused on the study although they may be imbalanced in terms of quality or quantity. The answer shows good discussion with reasonable understanding.	6–7
Balance of detail between strengths and weaknesses and both are focused on the study. Discussion is detailed with good understanding and clear expression.	8–10

Examples of possible discussion points:

Mann et al.

- *Strengths* Reliable because measured inter-rater reliability between coders **and it was high** (e.g. 0.99 for blinking, self-manipulations, illustrators etc.)
- Reliable because used objective measures such as pauses (r=0.55)
- Valid because it used real film of real criminals in high stakes situations
- Valid because observers were naïve so couldn't be biased and because the clips were corroborated
- Useful because it can avoid the police/courts being misled by believing certain behaviours tell us about lying
- *Weaknesess* Lacked reliability e.g. some measures were more subjective, such as such as pauses (r=0.55)
- Lacked reliability for some behaviours there were big individual differences (e.g. head movements, arm movements and gaze aversion, wth standard deviations around 10 or more)
- Lacked generalisibility because most criminals are not such high-stakes liars
- Not useful because there isn't a set of behaviours that we can be certain reveal lying.

Baron-Cohen et al.

- *strengths* able to collect objective data e.g. Eyes Test and AQ
- able to use statistical procedures enabling confirmation of patterns of eyes test responses in people with ASD.
- reliable measures because can be retested to demonstrate consistency e.g. IQ
- valid measures because can be compared to similar measures to show whether they are detecting appropriate characteristics e.g. ET revision
- Useful because it can lead us to help people with autism to look for emotions if they are trained what facial expressions mean.
- *weaknesses* responding to the eyes test involves decision making which could be subjective so assumptions about the validity and reliability of the quantitative data may be misplaced.
- quantitative data about one person's beliefs about the feelings of another person does not provide in-depth information about their understanding, if qualitative data had been collected they may have been able to provide more detailed insight into perception of emotional states e.g. eyes test cannot tell extent of empathy, only whether emotional state has been recognised
- some measures, such as IQ tests, may only measure a limited range of the intended variable, e.g. what about creative intelligence or specific intelligences?

Dement and Kleitman

- *strengths* able to collect objective data e.g. such as EEG and timing, which are reliable
- able to use statistical procedures leading to generalisations e.g. about typical patterns of such as being able to accurately estimate dream duration.
- collected both qualitative data revealed , which individual differences (e.g. in dream content) and quantitative data to make generalisations.
- *weaknesses* The participants were a limited sample so were likely to have all had similar cultural influence e.g. from the belief that 'we all need 8 hours sleep'.
- as a snapshot study the findings might not generalisible because situation is time limited so doesn't reflect real variability e.g. when we are stressed/ overworked/jetlagged etc. our sleep and dreaming patterns are disrupted
- not useful as it doesn't have any applications to the real world.