Cambridge Assessment International Education

Cambridge International AS & A Level

PSYCHOLOGY **9990/41**

Paper 4 Specialist Options: Application **May/June 2022**

1 hour 30 minutes

4466216065

You must answer on the enclosed answer booklet.

You will need: Answer booklet (enclosed)

INSTRUCTIONS

- Answer **four** questions in total:
  - Answer questions from **two** options.
    - Section A: answer **two** questions.
    - Section B: answer **one** question.
    - Section C: answer **one** question.
- Follow the instructions on the front cover of the answer booklet. If you need additional answer paper, ask the invigilator for a continuation booklet.

INFORMATION

- The total mark for this paper is 60.
- The number of marks for each question or part question is shown in brackets [ ].

This document has **8** pages. Any blank pages are indicated.

DC (RW) 303153/1
© UCLES 2022

[Turn over

stemup.co.uk

Section A

You must answer **two** questions from this section.
Answer all the parts for the two questions you choose.

Psychology and abnormality

1 In a case study by Glover, a 56-year-old married woman who had been shoplifting every day for 14 years attended treatment sessions. Her treatment used covert sensitisation with aversive imagery.

(a) Explain what is meant by the term 'aversive imagery'. [2]

(b) Suggest **two** strengths of the use of a face-to-face interview by a therapist treating a patient with kleptomania. [4]

(c) Suggest why biochemical treatments might **not** be effective when treating kleptomania. [4]

(d) Discuss whether generalisations can or cannot be made from this case study. You should consider both sides of the argument and include a conclusion. [5]

Psychology and consumer behaviour

2 Dayan and Bar-Hillel (2011) conducted research on how the position of a food item on a menu influenced what participants ordered to eat. They compared four versions of menu item order. They described their sample and sampling technique in Study 1 as:

- 240 participants
- from the Hebrew University, Jerusalem
- all were students
- aged 19–35 years
- 52% female, 48% male
- recruited individually from around the campus.

(a) Suggest which sampling technique could have been used in Study 1. [2]

(b) **(i)** Suggest **one** strength of the sample used in this study. [2]

(ii) Suggest **one** weakness of the sample used in this study. [2]

(c) **(i)** Identify **two** of the versions of menu item order used in Study 1. [2]

(ii) Explain why participants were randomly allocated to a version of menu item order. [2]

(d) Discuss the strengths and weaknesses of using a laboratory experiment to study menu item order. You should include a conclusion in your answer. [5]

© UCLES 2022

stemup.co.uk

Psychology and health

3 McKinstry and Wang (1991) investigated the way in which doctors' clothing influenced patients' confidence in doctors. Patients were shown photographs of different styles of doctors' clothing. Data was gathered by an interviewer asking closed questions in a structured interview.

(a) Explain what is meant by a 'structured interview'. [2]

(b) Give **one** strength and **one** weakness of asking closed questions as used in this study. [4]

(c) **(i)** Suggest how the quantitative data gathered for the styles in relation to confidence could be analysed. [2]

(ii) Suggest how the quantitative data gathered for the styles in relation to confidence could be shown on a graph. [2]

(d) Discuss the strengths and weaknesses of using photographs to gather information about a doctor's clothing. You should include a conclusion in your answer. [5]

Psychology and organisations

4 In just one year in the UK there were 21 deaths and 1372 injuries caused by chip pan fires. Cowpe (1989) reported on a safety campaign where two television advertisements were shown in ten regions of the UK between 1976 and 1982.

(a) Suggest the type of experiment used in this study. [2]

(b) Explain how **two** safety promotion strategies were used in the television advertisements for this campaign. [4]

(c) **(i)** Explain how the effectiveness of the campaign was measured. [2]

(ii) Suggest **one** strength of gathering data in this way. [2]

(d) Discuss the strengths and weaknesses of using television to promote safety in organisations. You should include a conclusion in your answer. [5]

stemup.co.uk

Section B

You must answer **one** question from this section.

Psychology and abnormality

5 **(a)** Design a study to investigate the effectiveness of biochemical treatments for impulse control disorder. [10]

(b) Explain the psychological and methodological evidence on which your study is based. [8]

Psychology and consumer behaviour

6 It is suggested that people have cognitive maps of where individual food items are located in retail environments, such as supermarkets.

(a) Design a study to investigate whether people use cognitive maps to locate individual food items in a supermarket. [10]

(b) Explain the psychological and methodological evidence on which your study is based. [8]

Psychology and health

7 **(a)** Design a study to investigate whether doctor-centred or patient-centred practitioners are more likely to make incorrect diagnoses. [10]

(b) Explain the psychological and methodological evidence on which your study is based. [8]

Psychology and organisations

8 **(a)** Design a study using a questionnaire to investigate which non-monetary reward is more effective for workers in an organisation. [10]

(b) Explain the psychological and methodological evidence on which your study is based. [8]

© UCLES 2022

stemup.co.uk

Section C

You must answer **one** question from this section.

Psychology and abnormality

9 *'Psychometric measures used to assess obsessive-compulsive disorder (OCD) provide therapists with no useful information.'*

To what extent do you agree with this statement? Use examples of research you have studied to support your answer. [12]

Psychology and consumer behaviour

10 *'Studies using an fMRI, such as those on pre-cognitive decisions, are of no value in understanding consumer behaviour.'*

To what extent do you agree with this statement? Use examples of research you have studied to support your answer. [12]

Psychology and health

11 *'A biochemical test is the only accurate way to know that a person has adhered to a request to take their medication.'*

To what extent do you agree with this statement? Use examples of research you have studied to support your answer. [12]

Psychology and organisations

12 *'There is no difference between Alderfer's ERG theory and Maslow's hierarchy of needs.'*

To what extent do you agree with this statement? Use examples of research you have studied to support your answer. [12]

© UCLES 2022

stemup.co.uk

BLANK PAGE

© UCLES 2022

stemup.co.uk

BLANK PAGE

© UCLES 2022

stemup.co.uk

BLANK PAGE

Permission to reproduce items where third-party owned material protected by copyright is included has been sought and cleared where possible. Every reasonable effort has been made by the publisher (UCLES) to trace copyright holders, but if any items requiring clearance have unwittingly been included, the publisher will be pleased to make amends at the earliest possible opportunity.

To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced online in the Cambridge Assessment International Education Copyright Acknowledgements Booklet. This is produced for each series of examinations and is freely available to download at www.cambridgeinternational.org after the live examination series.

Cambridge Assessment International Education is part of Cambridge Assessment. Cambridge Assessment is the brand name of the University of Cambridge Local Examinations Syndicate (UCLES), which is a department of the University of Cambridge.

© UCLES 2022

stemup.co.uk

Surname	StemUp Model Solution
Answer booklet	Write the question number in the left-hand margin at the start of each answer.

Question number

Answers for paper 9990/41

Section A

Psychology and abnormality

1(a) Aversive imagery is when the patient imagines a very unpleasant (aversive) scene and pairs it with the unwanted behaviour, so the behaviour becomes associated with negative feelings and is less likely to be repeated (a form of punishment). In Glover's case study the woman imagined feeling sick and vomiting as she approached an item in a supermarket, with other shoppers watching her, until she put the item back and left the store.

1(b) One strength is that the therapist can observe the patient's non-verbal behaviour, such as facial expression, hesitation or fidgeting, while she answers questions. This helps diagnosis and therapy, for example showing whether she is being honest about how often she shoplifts or which situations make the urge to steal strongest.

A second strength is that a face-to-face interview is more personal than a telephone interview. The therapist gives time to the patient and can build rapport and trust, so a patient with kleptomania, who may feel ashamed of stealing, is more likely to disclose details and to ask questions about her disorder and treatment.

1(c) There is no drug that specifically stops a person from stealing. Drugs can only

Question number	
	reduce general symptoms such as tension or anxiety; they do not remove the cause. Kleptomania involves a learned cycle of rising tension before the theft and pleasure or relief afterwards, so if this learned behaviour is not changed the patient is likely to relapse when the drug is stopped. Also, the patient may not adhere to the medication, for example if she does not see stealing as a medical problem, and long-term use of a drug may cause side effects or dependence, creating extra problems.
1(d)	Generalisations can be made because covert sensitisation is based on conditioning principles that apply to everyone, so the technique of pairing the urge to steal with an aversive image could be generalised to other patients with kleptomania. The same technique could also be generalised to other impulse control disorders, such as pyromania or pathological gambling, by changing the imagined scene. However, generalisations cannot be made about the specific details. This was one 56-year-old married woman who had stolen daily for 14 years, so her history, motivation and response to treatment are unique to her. Also, the particular aversive images of nausea and vomiting will not be unpleasant enough for everyone; other patients would need different images personal to them, so the exact procedure cannot be generalised. In conclusion, the general technique can be generalised but the individual procedure cannot, so the case study shows covert sensitisation is worth trying, but it should be tested on a larger sample with images tailored to each patient before it is assumed to work for everyone.
	Psychology and consumer behaviour
2(a)	An opportunity sample, because the participants were students recruited individually from around the campus, so they were simply the people who were

Question number	
	available and willing at the time the researchers approached them.
2(b)(i)	The sample had a nearly equal gender balance (52% female, 48% male), so menu choices were not biased towards one gender and the findings about item position can be generalised to both men and women.
2(b)(ii)	All participants were students from one university in Jerusalem aged only 19 to 35, so the findings may not generalise to older diners or people from other cultures, who may read and choose from menus differently.
2(c)(i)	The Baseline menu, which had the items in an arbitrary order, and the Mirror menu, which completely reversed the Baseline order within each food category.
2(c)(ii)	Random allocation gives each participant an equal chance of being in any of the four menu conditions, so participant variables such as food preferences or hunger are spread evenly. This avoids systematic bias, so any difference in what was ordered is due to the position of the items.
2(d)	A strength is control. The IV (the order of items on the menu) is manipulated and the DV (the item chosen) is measured while extraneous variables, such as the items, descriptions and prices, are kept the same, so cause and effect can be established. A second strength is that one variable, position on the menu, can be isolated and studied on its own, and the study is easy to replicate to check reliability. A weakness is low ecological validity. Real ordering happens in restaurants with waiters, friends, smells, specials and real money, so choices on a menu in a study may not reflect what people really order. Also, participants know they are in a study, so they may respond to demand characteristics, for example avoiding the

Question number	
	first item if they guess the aim, which lowers validity. In conclusion, a laboratory experiment is the best way to show that menu position alone can change choices, but because real dining involves many other variables, its findings should be checked with a field experiment in a real restaurant before being applied by menu designers.
	Psychology and health
3(a)	A structured interview is one where the interviewer asks every participant the same pre-set questions in the same order, with no extra questions added. For example, McKinstry and Wang asked every patient which doctor they would feel happiest about seeing for the first time.
3(b)	A strength is that closed questions produce quantitative data, because every patient answers on the same rating scale. This means the confidence scores for the different styles of clothing, such as a formal suit or casual clothes, can be compared statistically, and the study can easily be replicated. A weakness is that patients cannot explain the reason for their answer. A patient might prefer the doctor in formal clothes because of the doctor's age or expression rather than the clothing, but closed questions give no opportunity to add this information, so validity may be lower.
3(c)(i)	Calculate the mean confidence score for each style of clothing (for example white coat, formal suit, casual), then compare the means to see which style gives the most confidence, and use the range to show how varied the patients' ratings were.
3(c)(ii)	A bar chart, with a separate bar for each style of doctor's clothing on the x-axis and the mean confidence score on the y-axis, so the style with the tallest bar

Question number

gives patients the most confidence.

3(d) A strength is standardisation: in a photograph the doctor's expression, pose and background stay exactly the same, so the only thing that changes between photographs is the clothing, which increases validity. A second strength is that the study can be replicated with the same photographs, and they can be shown to patients in many different practices, not just one place.

A weakness is that photographs are static, while real doctors move, speak and show a manner and tone of voice. Patients' confidence in a real consultation depends on these too, so judging clothing from a photograph has low ecological validity. Also, the people shown in the photographs must give consent for their images to be used, and patients may focus on the person's face or age rather than the clothing.

In conclusion, photographs are the most controlled way to isolate the effect of clothing, so they are useful for a first study, but the findings should be checked with real doctors in real consultations before practices make rules about what doctors wear.

Psychology and organisations

4(a) A field experiment, because the IV (whether the television advertisements were shown in a region) was manipulated by the researchers, and it took place in a real-life setting, people's homes in ten regions of the UK, with the DV being the number of chip pan fires.

4(b) One strategy was prevention: the advertisements showed people how to avoid a chip pan fire starting, for example by not overfilling the pan with oil and never leaving it unattended while it is heating.

Question number	
	A second strategy was containment: the advertisements showed the correct procedure if a fire does start, such as turning off the heat, covering the pan with a damp tea towel and leaving it to cool, and showed that throwing water on it or carrying the pan outside is wrong. Showing a real fire and people's shocked reactions also used fear arousal to make the message memorable.
4(c)(i)	By the number of chip pan fires recorded in official fire brigade statistics in each region, comparing the number of fires before the advertisements with the number during and in the months after they were shown.
4(c)(ii)	The data is objective: a fire brigade call-out either happened or it did not, so it cannot be biased by social desirability, unlike asking people whether they had a chip pan fire. It is also quantitative, so the number of fires in different regions can be compared.
4(d)	A strength is reach: large numbers of the population, including most workers, watch television, so a safety message reaches far more people than a poster in one workplace. A second strength is impact: television can show a real hazard, such as a fire, and model the correct behaviour, so it can arouse fear and show people exactly what to do, which makes the message memorable. It can also be targeted at one region or channel. A weakness is that some people do not have a television, watch streaming services without adverts or skip the adverts, so they never see the message. Also, television messages are general, so they cannot cover the specific hazards of one organisation, such as a particular machine, and the organisation cannot check whether workers understood or followed the advice. In conclusion, television is effective for raising general safety awareness, as the fall in chip pan fires suggests, but within an organisation it should be backed by

Question number

on-site methods such as training or a token economy for safe behaviour (Fox et al., 1987), which target its own hazards.

Section B

Psychology and abnormality

5(a) I would carry out a randomised, double-blind, placebo-controlled experiment on pathological gambling, an impulse control disorder, testing the drug nalmefene, an opioid antagonist.

Sample: 80 adults aged 18 to 65 diagnosed with pathological gambling by a psychiatrist using DSM criteria, recruited as a volunteer sample through advertisements in gambling-support clinics. People with another disorder, such as bipolar disorder, or already taking medication would be excluded.

Design and IV: an independent measures design. Participants would be randomly allocated to two conditions: nalmefene tablets or identical-looking placebo tablets, taken daily for 16 weeks. Neither the participants nor the psychiatrists rating them would know who had which tablet (double blind).

DV: the urge to gamble, measured with the PG-YBOCS, a version of the Yale-Brown scale for gambling, rated by a psychiatrist at the start and every four weeks. Participants would also keep a weekly diary of the number of gambling episodes and the money spent, giving a second, quantitative measure.

Controls: same dose, same timing, same rating scale and same interview procedure for everyone; no other treatment during the study.

Reliability and validity: two psychiatrists would rate each patient independently to check inter-rater reliability. Validity is improved by the placebo group and double blind, which control expectations and experimenter bias.

Ethics: fully informed consent, the right to withdraw at any time, monitoring of

Question number

side effects, confidential data, and a full debrief, with the placebo group offered the real treatment at the end.

Analysis: the mean reduction in PG-YBOCS score would be compared for the two groups using a Mann-Whitney U test, and shown in a bar chart.

5(b) Psychologically, the study is based on the biochemical explanation of impulse control disorders. Gambling gives a "high" through opioids and dopamine in the brain's reward system. Nalmefene is an opioid antagonist which blocks opioid receptors, so it should reduce the pleasure of gambling and therefore the urge to gamble. Grant et al. (2008) tested nalmefene on pathological gamblers: participants were assessed to make sure they were suitable, then randomly allocated to nalmefene or a placebo, and the urge to gamble was measured with the PG-YBOCS. My study copies this design so the results can be compared with theirs.

Methodologically, a laboratory-style clinical experiment allows the IV (drug or placebo) to be manipulated and other variables controlled, so a cause and effect conclusion about the drug can be drawn. Independent measures is needed because a participant cannot take the drug and not take it, and random allocation spreads participant variables, such as severity, evenly. The placebo controls for the expectation of getting better, and the double-blind procedure stops participants and raters showing demand characteristics or bias. Using a standardised clinician-rated scale gives quantitative, reliable data, while the gambling diary adds a measure of real behaviour, improving validity. Ethical guidelines are followed because the participants are vulnerable patients who must be protected from harm.

Psychology and consumer behaviour

6(a) I would carry out a quasi-experiment in a large supermarket, comparing regular

Question number

shoppers with people who have never shopped there.

Sample: 30 regular shoppers (who shop there at least once a week) and 30 new shoppers, aged 18 to 65, selected by opportunity sampling at the entrance, with permission from the store manager.

IV: familiarity with the store (regular or new). This is naturally occurring, so the design is independent measures.

Procedure and DVs: before entering, each participant is given a blank outline of the store and asked to draw a sketch map showing where 10 named items are, such as milk, rice, bread and toothpaste. Score: the number of items placed in the correct aisle (out of 10). Each participant is then given the same list and asked to collect the items. A researcher follows them (a non-participant, overt observation) and records the time taken and the number of wrong aisles entered, marking their route on a store plan.

Controls: the same 10 items for everyone, testing on weekday mornings when the store is quiet, and no changes to the store layout during the study.

Reliability: two researchers would score each sketch map independently to check inter-rater reliability.

Ethics: informed consent, the right to withdraw, data kept anonymous, and participants keep any items they want to buy as normal.

Analysis: mean sketch-map scores and mean search times for each group compared with a Mann-Whitney U test; if regular shoppers draw more accurate maps and find items faster, this suggests they use a cognitive map.

6(b) Psychologically, a cognitive map is a mental representation of a spatial layout that a person uses to find their way around. MacKay and Olshavsky (1975) studied consumers' cognitive maps of retail locations, and one way they measured them was by asking people to draw sketch maps, which is why my study uses a sketch-map task. If people have a cognitive map of a supermarket, it should be

Question number

more accurate the more often they shop there, so familiarity is the IV. Gil et al. (2009) used CCTV to record shopper movement patterns, showing that routes through a store can be recorded, which supports observing and plotting each participant's route.

Methodologically, a quasi-experiment is used because familiarity cannot be manipulated; it already exists. Independent measures is needed as a person cannot be both a regular and a new shopper. Using two DVs, a sketch map (what people know) and search time (what people do), is a form of triangulation, so the conclusion is more valid than with one measure. Carrying out the study in a real supermarket gives high ecological validity, while the controls (same items, same time of day) reduce extraneous variables. Quantitative scores allow statistical comparison, and checking inter-rater reliability makes the scoring of maps consistent. Opportunity sampling is practical at a store entrance, and consent protects participants.

Psychology and health

7(a) I would use a field study with simulated patients (trained actors) visiting family doctors, combined with a structured observation of the consultations.

Sample: 40 family doctors (GPs) from different practices, recruited as volunteers. They agree that over the next six months some of their patients may be actors, but do not know which.

Step 1, classifying doctors: with consent, 10 of each doctor's ordinary consultations are audio-recorded. Two trained observers code them using behavioural categories, such as doctor-centred (closed questions, focusing on the first problem mentioned, giving instructions) and patient-centred (open questions, listening, asking the patient's views). Each doctor is classed as doctor-centred or patient-centred by their main style. This gives the IV, so it is a quasi-experiment with independent

Question number	
	groups. Step 2, measuring errors: each doctor is visited by four actors with scripted symptoms: two have a known condition (such as the symptoms of a migraine) and two are healthy but describe a minor worry. DV: the number of incorrect diagnoses out of four, counted as false positives (diagnosing an illness in a healthy patient) and false negatives (missing the real condition). Controls: all actors use the same scripts, the same symptoms and the same appointment length. Reliability: two observers code the recordings, checking inter-rater reliability. Ethics: consent from doctors, patients consent to being recorded, confidentiality of doctors' results, and a debrief explaining which visits were actors. Analysis: percentage of incorrect diagnoses for each style compared with a chi-squared test.
7(b)	Psychologically, the study is based on Byrne and Long (1976), who analysed over 2000 tape-recorded GP consultations and identified a range of styles from doctor-centred, where the doctor controls the consultation, to patient-centred, where the patient is involved. My coding categories are based on their behaviours. Savage and Armstrong (1990) also compared a directive (doctor-centred) and a sharing (patient-centred) style. A doctor-centred doctor who asks only closed questions may miss information and so make more diagnostic errors. The study measures both types of error: Rosenhan (1973) showed that doctors are more inclined to call a healthy person sick (a false positive, which he called a type 2 error) than to call a sick person healthy, which is why healthy actors are included. Methodologically, a structured observation with clear behavioural categories gives quantitative data and can be checked for inter-rater reliability. Using simulated patients means the correct diagnosis is known, so errors can be measured objectively, and the same scripts act as a control, making the visits standardised.

Question number

Real consultations give high ecological validity, and doctors not knowing which patients are actors reduces demand characteristics. A quasi-experiment is used because style cannot be manipulated. Ethics are addressed with prior consent and a full debrief.

Psychology and organisations

8(a) I would use a questionnaire given to employees of a large supermarket chain.

Sample: 200 employees chosen by stratified sampling, so the proportions of shop-floor staff, supervisors and managers, and of men and women, match the organisation.

Type and setting: an anonymous online questionnaire completed during work time in a private room, taking about 15 minutes.

Questions: Section 1 has closed Likert-scale items for five non-monetary rewards: praise, respect, recognition, empowerment and a sense of belonging, for example "Being praised by my manager makes me work harder" and "Being given the power to make decisions in my job motivates me", rated from 1 (strongly disagree) to 5 (strongly agree). Some items are reverse worded, such as "Being thanked in public makes no difference to my effort", to reduce response bias. Section 2 asks workers to rank the five rewards from 1 (most motivating) to 5. Section 3 has an open question: "Describe a time when a reward that was not money made you work harder."

Scoring and analysis: the mean Likert score and the median rank are calculated for each reward and shown in a bar chart; the reward with the highest mean is the most effective. Open answers are analysed for themes.

Reliability: the questionnaire is given again two weeks later to check test-retest reliability.

Validity: a pilot study checks the questions are clear.

Question number	
	Ethics: informed consent, anonymity, the right to withdraw, and managers never see individual answers.
8(b)	Psychologically, the study is based on the types of non-monetary reward: praise (being told you have done well), respect (being valued as a person), recognition (achievement being publicly acknowledged, such as employee of the month), empowerment (being given responsibility and decision-making power) and a sense of belonging (feeling part of the team). These link to need theories: Maslow (1970) argues that people are motivated by belonging and esteem needs, which belonging, praise and recognition meet, and Herzberg's two-factor theory treats recognition and responsibility as motivators. The study tests which of these rewards is actually most effective. Methodologically, a questionnaire allows a large sample to be asked quickly, and stratified sampling makes it representative of the whole organisation. Closed Likert questions give quantitative data so the rewards can be compared with means, while ranking forces a clear choice. The open question gives qualitative data explaining why a reward works, adding validity. Anonymity reduces social desirability bias, which is likely when workers are asked about their managers, and reverse-worded items reduce response bias. Test-retest checks reliability, and a pilot study improves validity by removing unclear questions. Ethical guidelines protect workers from any consequences at work.
	Section C
	Psychology and abnormality
9	Psychometric measures are standardised tests used to assess a disorder. Two used for OCD are the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) and the

Question number

Maudsley Obsessive-Compulsive Inventory (MOCI). The Y-BOCS is clinician-rated: the therapist asks about obsessions and compulsions separately, rating items such as time spent, distress and resistance from 0 to 4, giving a total out of 40. The MOCI is a self-report of 30 true or false statements with subscales such as checking, washing, slowness and doubting.

There is strong evidence against the statement. These measures give the severity of the disorder as a number, so a therapist can decide whether treatment is needed and which treatment suits the patient. They also show the specific features of a person's OCD: the MOCI subscales show whether the problem is mainly washing or checking, so exposure and response prevention can target the right behaviour. Because the data is quantitative, scores can be compared with other people with the disorder and measured again during therapy to track progress. For example, in Lehmkuhl et al.'s (2008) case study of Jason, a 12-year-old with autism and OCD, a children's version of the Y-BOCS was used to show that his symptoms fell after CBT.

However, there are limitations that partly support the statement. Self-reports such as the MOCI rely on honesty, and people with OCD may hide embarrassing obsessions because of social desirability. Fixed answers such as true or false, or a rating scale with a neutral option, may not represent what the person really thinks. The measures are nomothetic: they assume people with OCD share similar symptoms, so an unusual obsession may be missed. A clinical interview can reveal more, such as the meaning of the obsession or what triggered it, which gives qualitative detail a score cannot.

In conclusion, I largely disagree with the statement. Psychometric measures clearly provide useful information about severity, type of symptom and progress, which is why therapists use them. Their weaknesses mean they should not be used alone, but combined with a clinical interview they give therapists a valid and reliable picture of the patient.

Question number

Psychology and consumer behaviour

10 fMRI measures brain activity by detecting changes in blood flow while a person does a task. Knutson et al. (2007) used fMRI to study pre-cognitive decisions about buying. Participants in the scanner were shown a product, then its price, and then chose whether to buy it. Activation in the nucleus accumbens when the product was shown predicted that the person would buy it, while activation in the insula when the price was seen as too high predicted that they would not buy. Activity in the medial prefrontal cortex was linked to the price. This suggests the decision starts in the brain before the person is aware of it.

There is evidence that such studies are of value. fMRI is scientific equipment producing objective, quantitative data that is not affected by social desirability, unlike asking consumers why they bought something. Studying how the mind makes decisions is always useful, and Knutson's findings show that emotion (anticipated pleasure) and the "pain" of paying both affect purchasing. This data could be useful to sellers: if they know which features activate reward areas or what price triggers the insula, they can design and price products people are more likely to buy.

On the other hand, there are reasons to agree with the statement. Lying still in a noisy scanner looking at pictures has low ecological validity compared with shopping in a store with friends, music and time pressure. Samples are small and fMRI is expensive, so findings are hard to generalise. The data is correlational: brain activity is linked to a choice but does not explain why. It is also reductionist, ignoring social and cultural factors. Finally, a decision made before awareness may be irrelevant if the person then acts on it; simply observing what people buy is also factual and is cheaper.

In conclusion, I disagree that fMRI studies have no value, as they give unique

Question number

objective evidence about how buying decisions are made. However, their value is limited by low ecological validity and cost, so they are most useful when combined with real-world methods such as observation of actual purchases.

Psychology and health

11 Adherence means following medical advice, such as taking medication as prescribed. A biochemical test, such as a blood or urine test, detects whether the drug, or a marker added to it, is in the body.

There is support for the statement. Biochemical tests are objective and give quantitative data that the patient cannot manipulate: if the drug is not in the blood, it has not been taken. Roth and Caron (1978) used a biochemical measure of whether patients with ulcers took their antacid medication and found that both patients and doctors overestimated adherence compared with this objective measure. This shows that self-report is not accurate, as patients may forget or give socially desirable answers, such as in self-report questionnaires (for example Riekert and Drotar, 1999).

However, there are strong arguments against the statement. A biochemical test may not be valid: a patient could take the drug only on the day before the test, and the test shows that the drug was taken recently, not that it was taken correctly every day. Tests are invasive (blood samples), costly and time-consuming, so in practice they are rarely done. Also, there are other objective measures. Chung and Naya (2000) used an electronic pill counter (TrackCap), which records each time the bottle is opened, giving a detailed record of when doses were taken, which a single blood test cannot. Repeat prescriptions, as used by Sherman et al. (2000), show whether a patient collects their medication on time, without any invasive procedure. However, these also have limits: opening a bottle or collecting a prescription does not prove the pill was swallowed.

Question number

In conclusion, I disagree that a biochemical test is the only accurate way. It is the most direct measure of whether a drug is in the body, but it only shows one moment in time. The most accurate picture comes from combining methods, for example a biochemical test with electronic pill counting, so each method's weaknesses are covered by the other.

Psychology and organisations

12 Maslow's (1970) hierarchy of needs is a need theory of motivation in which lower needs (physiological, safety) must be met before higher needs (belonging, esteem, self-actualisation), with later versions adding cognitive, aesthetic and transcendence needs to make eight levels. Alderfer's (1972) ERG theory has three needs: existence, relatedness and growth.

There is some support for the statement. Both are need theories, and both say workers are motivated by unmet needs. ERG theory is based on Maslow's needs: existence combines physiological and safety needs, relatedness covers belonging and the social part of esteem, and growth covers self-esteem and self-actualisation. Both recognise the importance of the same basic needs, such as pay and safe conditions, and both suggest the same practical advice for managers, such as meeting basic needs first and then giving opportunities for development.

However, there are important differences, so I mostly disagree. Maslow has eight (or five) levels while Alderfer has three, and combining two needs into one may ignore features of each. Maslow's hierarchy is fixed: a person moves up only when a lower need is satisfied. Alderfer says several needs can motivate at the same time, so a worker may want growth while existence needs are not fully met. Alderfer also added the frustration-regression principle: if a growth need is frustrated, the worker regresses and puts more effort into relatedness, for example socialising more at work, which Maslow does not include. Maslow's later needs,

Question number

such as aesthetic and transcendence needs, do not appear in ERG theory. ERG theory was developed from studies of workers in organisations and is more flexible, so it explains individual differences, for example why people in different cultures rank needs differently, better than Maslow's fixed order.

In conclusion, I disagree with the statement. ERG theory grew out of Maslow's hierarchy and they share the same basic types of need, but ERG theory differs in its number of needs, its flexible order and the frustration-regression principle, making it a different and more applicable theory for the workplace.

Cambridge Assessment International Education

Cambridge International AS & A Level

PSYCHOLOGY **9990/42**

Paper 4 Specialist Options: Application **May/June 2022**

1 hour 30 minutes

You must answer on the enclosed answer booklet.

You will need: Answer booklet (enclosed)

INSTRUCTIONS

- Answer **four** questions in total:
  - Answer questions from **two** options.
    - Section A: answer **two** questions.
    - Section B: answer **one** question.
    - Section C: answer **one** question.
- Follow the instructions on the front cover of the answer booklet. If you need additional answer paper, ask the invigilator for a continuation booklet.

INFORMATION

- The total mark for this paper is 60.
- The number of marks for each question or part question is shown in brackets [ ].

This document has **8** pages. Any blank pages are indicated.

© UCLES 2022

[Turn over

stemup.co.uk

Section A

You must answer **two** questions from this section.
Answer all the parts for the two questions you choose.

Psychology and abnormality

1 Schizophrenia can be treated with medical techniques and it can be managed with psychological techniques. Medical treatments include the use of electro-convulsive therapy (ECT) and drugs such as antipsychotics and atypical antipsychotics. Schizophrenia can be managed psychologically using token economy and cognitive-behavioural therapy (CBT).

(a) Suggest **one** difference between medical and psychological techniques for treating and managing schizophrenia. [2]

(b) Explain **two** differences between antipsychotic drugs and atypical antipsychotic drugs. [4]

(c) Explain how ECT is used to treat schizophrenia. [4]

(d) Discuss the strengths and weaknesses of psychological techniques to manage schizophrenia. You should include a conclusion in your answer. [5]

Psychology and consumer behaviour

2 Porublev et al. (2009) used interviews and observations in their investigation into gift-wrapping. In one semi-structured interview, a participant stated: 'What's the point of giving a gift if you are not going to wrap it?'

(a) Explain **one** way in which a gift can be wrapped, when it is presented to a receiver. [2]

(b) Suggest **two** reasons why a semi-structured interview was used in this study. [4]

(c) Suggest **two** features of the type of observation used in this study. [4]

(d) Discuss the strengths and weaknesses of using observations to assess gift-wrapping preference. You should include a conclusion in your answer. [5]

© UCLES 2022

stemup.co.uk

Psychology and health

3 Biofeedback can be used for the management of stress. Budzynski et al. (1969) used biofeedback to reduce tension headaches. Three conditions were used:

- variable pitch (correct feedback)
- constant low tone (irrelevant feedback)
- silent (no feedback).

The results of the study are shown in Fig. 3.1.

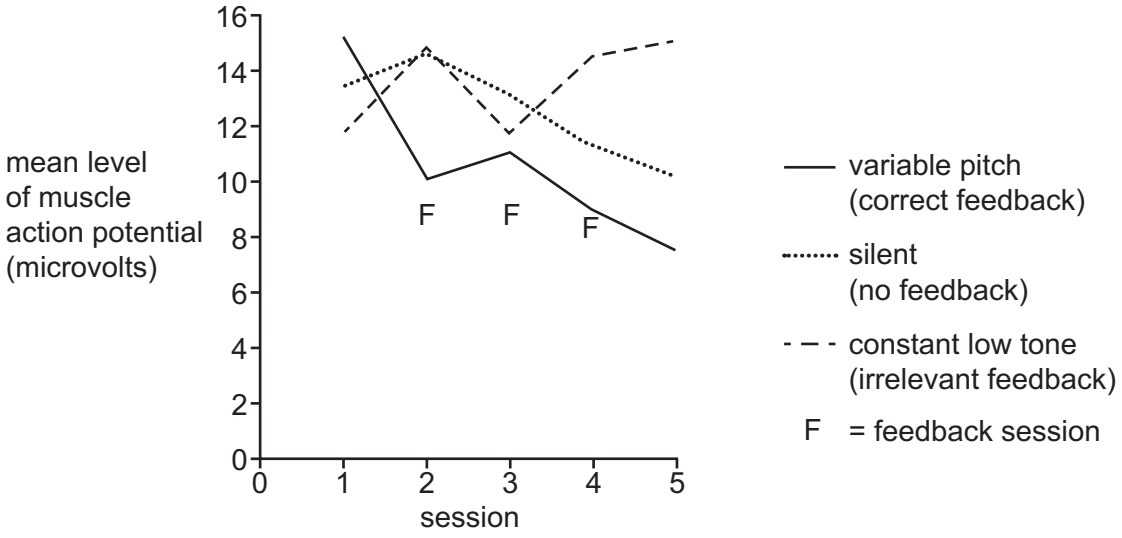


Fig. 3.1

(a) Outline what is meant by 'biofeedback', using an example. [2]

(b) Give **two** findings from the data in Fig. 3.1. [4]

(c) Suggest **two** techniques that can be used for the management of stress, other than biofeedback. [4]

(d) Discuss the strengths and weaknesses of using biofeedback for the management of stress. You should include a conclusion in your answer. [5]

Psychology and organisations

4 Job design involves matching work processes to improve a worker's interest in the task and to improve productivity. It can involve job rotation, job enrichment and job enlargement.

(a) Explain what is meant by the term 'job rotation'. [2]

(b) Suggest **one** similarity and **one** difference between job enrichment and job enlargement. [4]

(c) **(i)** Suggest **one** way in which worker satisfaction can be measured using a psychometric test. [2]

(ii) Suggest **one** way in which worker satisfaction can be measured, other than using a self-report. [2]

(d) Discuss the strengths and weaknesses of measuring job satisfaction using an interview. You should include a conclusion in your answer. [5]

stemup.co.uk

Section B

You must answer **one** question from this section.

Psychology and abnormality

5 **(a)** Design a study to investigate the effectiveness of chemical/drug treatments (MAOIs, SSRIs) for depression. [10]

(b) Explain the psychological and methodological evidence on which your study is based. [8]

Psychology and consumer behaviour

6 **(a)** Design an experiment to investigate age differences in brand recognition. [10]

(b) Explain the psychological and methodological evidence on which your experiment is based. [8]

Psychology and health

7 **(a)** Design a study using a questionnaire to investigate gender differences in hypochondriasis. [10]

(b) Explain the psychological and methodological evidence on which your study is based. [8]

Psychology and organisations

8 **(a)** Design a study to investigate absenteeism on different days of the week in a food production factory. [10]

(b) Explain the psychological and methodological evidence on which your study is based. [8]

© UCLES 2022

stemup.co.uk

Section C

You must answer **one** question from this section.

Psychology and abnormality

9 *'The nature versus nurture debate is irrelevant when explaining depression.'*

To what extent do you agree with this statement? Use examples of research you have studied to support your answer. [12]

Psychology and consumer behaviour

10 *'In all advertising media, there is only **one** effective way to advertise a product, for example television.'*

To what extent do you agree with this statement? Use examples of research you have studied to support your answer. [12]

Psychology and health

11 *'Psychological measures of stress are more valid and reliable than physiological measures.'*

To what extent do you agree with this statement? Use examples of research you have studied to support your answer. [12]

Psychology and organisations

12 *'Belbin's team role types are the perfect way to categorise people into team roles.'*

To what extent do you agree with this statement? Use examples of research you have studied to support your answer. [12]

© UCLES 2022

stemup.co.uk

BLANK PAGE

© UCLES 2022

stemup.co.uk

BLANK PAGE

© UCLES 2022

stemup.co.uk

BLANK PAGE

Permission to reproduce items where third-party owned material protected by copyright is included has been sought and cleared where possible. Every reasonable effort has been made by the publisher (UCLES) to trace copyright holders, but if any items requiring clearance have unwittingly been included, the publisher will be pleased to make amends at the earliest possible opportunity.

To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced online in the Cambridge Assessment International Education Copyright Acknowledgements Booklet. This is produced for each series of examinations and is freely available to download at www.cambridgeinternational.org after the live examination series.

Cambridge Assessment International Education is part of Cambridge Assessment. Cambridge Assessment is the brand name of the University of Cambridge Local Examinations Syndicate (UCLES), which is a department of the University of Cambridge.

© UCLES 2022

stemup.co.uk

Surname	StemUp Model Solution
Answer booklet	Write the question number in the left-hand margin at the start of each answer.

Question number

Answers for paper 9990/42

Section A

Psychology and abnormality

1(a) Medical techniques change the person's physiology, for example antipsychotic drugs block dopamine receptors, whereas psychological techniques change the person's thinking or behaviour, for example CBT challenges delusional beliefs. Also, in medical treatment the patient is passive, while in psychological treatment the patient must be active.

1(b) First, typical antipsychotics are first-generation drugs introduced in the 1950s, such as chlorpromazine, whereas atypical antipsychotics are second-generation drugs that came into wide use in the 1990s, such as clozapine and risperidone.

Second, typical antipsychotics are more likely to cause movement side effects such as extrapyramidal symptoms and tardive dyskinesia (abnormal movements of the face, tongue and mouth), because they block dopamine receptors strongly and for a long time. Atypical antipsychotics block dopamine receptors more briefly, so these side effects are less likely, although weight gain and diabetes are more likely.

1(c) The patient is given a general anaesthetic and a muscle relaxant. Electrodes are placed on the head, either on both sides (bilateral) or on one side (unilateral), and a brief electric current is passed through the brain. This causes a seizure, but

Question number

because of the muscle relaxant the body only twitches slightly. The patient is unconscious during this, then wakes a few minutes later and recovers. It is usually given as a course of several sessions over a few weeks.

1(d) A strength is that the patient is active in their treatment rather than passive, so they learn to cope with their symptoms. For example, in CBT patients learn to test and challenge their delusions, and Sensky et al. (2000) found that patients given CBT continued to improve at follow-up. A second strength is that no drugs are involved, so there are no side effects such as tardive dyskinesia, and techniques like CBT can be applied anywhere once learned. Paul and Lentz (1977) also found token economy helped long-stay patients gain everyday skills.

A weakness is that psychological techniques need a trained therapist and many sessions, so they are more costly and time-consuming than giving a drug. Also, some patients prefer to be passive, as swallowing a pill is easy, and patients in an acute psychotic state may not be able to engage. Psychological techniques also do not change the dopamine activity linked to schizophrenia, which drugs do.

In conclusion, psychological techniques are valuable because they give patients lasting coping skills without side effects, but they work best alongside medication, which first reduces the symptoms enough for the patient to take part in therapy.

Psychology and consumer behaviour

2(a) A gift can be wrapped in a traditional manner, using patterned gift paper with ribbon or a bow, so that it meets the receiver's expectations and is clearly recognisable as a gift as soon as it is presented.

2(b) First, a semi-structured interview has some fixed questions, so every participant is asked the same core questions, such as "Do you prefer to receive gifts that are

Question number	
	wrapped or unwrapped?". This means answers can be compared and the number preferring wrapped gifts can be counted as quantitative data. Second, it allows the interviewer to ask extra, open follow-up questions, such as "Why do you prefer wrapped gifts?". This lets participants expand on their answers, giving rich qualitative data such as the comment "What's the point of giving a gift if you are not going to wrap it?", which explains people's reasons.
2(c)	One feature is that it was a naturalistic observation, because people were observed in a real setting, at a gift-wrapping stall, while they were actually buying and wrapping gifts, with nothing manipulated. A second feature is that it was a non-participant observation, because the observer did not wrap gifts or take part, but watched and recorded whether people paid to have their gifts wrapped.
2(d)	A strength is that observing people at a real gift-wrapping stall is naturalistic, so the behaviour is genuine: people are actually choosing whether to pay for wrapping, which is more valid than saying what they would do. Another strength is that a structured observation means observers know exactly what they are looking for, such as whether a gift was wrapped, and with two observers inter-rater reliability can be checked. A weakness is that observation records only behaviour, not reasons, so we do not know why a person chose wrapping, for example whether it was for the receiver or for convenience, and what was said between giver and receiver is not recorded. Also, if the observation is covert, people have not consented, and an observer may miss behaviour if a person moves out of view. In conclusion, observation gives a valid measure of what people really do about gift-wrapping, but it should be combined with interviews, as Porublev et al. did, to explain the reasons behind the preference.

Question number

Psychology and health

3(a) Biofeedback is a technique in which a person is given information about a physiological process so they can learn to control it psychologically. For example, electrodes measure muscle tension in the forehead and turn it into a tone whose pitch rises with tension; by relaxing, the person learns to lower the tone.

3(b) First, the variable pitch (correct feedback) group started with the highest mean muscle action potential at session 1, about 15 microvolts, but ended as the lowest at session 5, about 7.5 microvolts, so their muscle tension roughly halved. Second, the constant low tone (irrelevant feedback) group started as the lowest at session 1, about 12 microvolts, but ended as the highest at session 5, about 15 microvolts, so their tension increased.

3(c) One technique is stress inoculation training (Meichenbaum, 1985), a cognitive-behavioural therapy with three phases: conceptualisation, skill acquisition and rehearsal, and application and follow-through, so the person learns coping skills before meeting stressors.
A second technique is imagery and relaxation, as used by Bridge et al. (1988) with women having radiotherapy for breast cancer, where patients learned to relax their muscles and imagine a peaceful scene, which reduced their stress.

3(d) A strength is that the person controls the treatment themselves, so once learned it can be applied wherever they are, for example relaxing their forehead muscles when a tension headache starts. A second strength is that it gives immediate visual or auditory feedback on their level of stress, which is rewarding and helps learning, as shown by the correct feedback group in Budzynski et al. (1969)

Question number	
	reducing their muscle tension. There are also no drug side effects. A weakness is that biofeedback needs time, equipment and a trained therapist before people become skilled, so it is not a quick solution like taking a drug. Also, some people are not good at it or prefer to be passive and have a medically qualified person treat them, and it reduces the physical symptoms without dealing with the cause of the stress, such as problems at work. In conclusion, biofeedback is an effective, side-effect-free way to manage the physical symptoms of stress for motivated people, but it is best combined with a technique such as stress inoculation training that also tackles the causes.
	Psychology and organisations
4(a)	Job rotation is when a worker moves between different tasks at regular intervals while staying at the same level of skill and responsibility, for example a factory worker who packs on Monday and labels on Tuesday. It adds variety and reduces boredom but is unlikely to increase job satisfaction.
4(b)	A similarity is that both add more tasks to a worker's job, so both increase the variety of the work and reduce boredom, which may increase satisfaction. A difference is that job enrichment is a vertical extension: the worker gains more responsibility and new, higher-level skills, such as planning their own schedule or checking quality. Job enlargement is a horizontal extension: the worker does more tasks at the same skill level and the same responsibility, such as assembling and also packing a product.
4(c)(i)	Use the Minnesota Satisfaction Questionnaire (Weiss et al., 1967). Workers rate satisfaction with aspects of their job, such as supervision and working conditions, on a 5-point scale from very dissatisfied to very satisfied, and the scores are

Question number	
	added to give a satisfaction score.
4(c)(ii)	Record each worker's absenteeism and lateness from company records over several months; frequent absence or lateness suggests low job satisfaction, while good attendance suggests high satisfaction.
4(d)	A strength is that an interview allows workers to explain their views on their job in full detail, giving rich qualitative data about why they are satisfied or not, for example with their manager. A second strength is flexibility: an unstructured interview allows free discussion, while a structured one ensures management covers essential questions, and it can be held in a private, safe environment and recorded as evidence. A weakness is that workers may not tell the truth, hiding or modifying information because of social desirability or fear that management will punish them for criticism. Also, some workers may be reluctant to be interviewed at all, or worry they will be seen as weak or inadequate if they admit dissatisfaction, which reduces validity. In conclusion, interviews give the most detailed picture of job satisfaction, but their validity depends on trust, so they are best carried out by an independent interviewer with guaranteed confidentiality, alongside an anonymous questionnaire.
	Section B
	Psychology and abnormality
5(a)	I would carry out a randomised, double-blind, placebo-controlled experiment in hospital outpatient clinics. Sample: 90 adults aged 18 to 60 diagnosed with major depression by a psychiatrist

Question number	
	using DSM criteria, not currently taking medication, recruited as a volunteer sample through their doctors.
	IV and design: independent measures with three conditions: an SSRI (fluoxetine), an MAOI (phenelzine) or a placebo, all as identical capsules, taken daily for 12 weeks. Participants are randomly allocated, 30 per group. All participants receive the same dietary advice, as is needed for MAOIs, so the groups cannot be told apart. Neither the participants nor the psychiatrists rating them know which capsule each person has.
	DV: severity of depression measured with the Beck Depression Inventory (BDI), 21 items each scored 0 to 3, at the start and at weeks 4, 8 and 12, with a follow-up at 6 months. A psychiatrist also rates each patient in a structured interview as a second measure.
	Controls: same dosing schedule, same appointment procedure, no other therapy during the study.
	Reliability and validity: the BDI is standardised; two psychiatrists rate each interview to check inter-rater reliability; the placebo and double blind control expectation effects and bias.
	Ethics: fully informed consent, the right to withdraw, close monitoring for side effects and suicide risk, with anyone who gets worse removed and treated, confidentiality, and the placebo group offered an effective drug afterwards.
	Analysis: the mean fall in BDI score for each group, shown on a line graph over time and compared with a Kruskal-Wallis test.
5(b)	Psychologically, the study is based on the biological explanation that depression is linked to low levels of monoamine neurotransmitters such as serotonin and noradrenaline. MAOIs inhibit the enzyme monoamine oxidase, which normally breaks down noradrenaline, serotonin and dopamine, so these stay at higher levels and depression should reduce. SSRIs block the reuptake of serotonin into the

Question number

presynaptic neuron, so more serotonin remains in the synapse. Comparing the two tests which mechanism is more effective. The BDI (Beck et al., 1961) is used as it is a well-established measure of the cognitive and physical symptoms of depression. A placebo group is essential because Kirsch et al. (2008) found that much of the improvement with antidepressants is also seen with placebos. Methodologically, an experiment allows the IV (type of drug) to be manipulated with other variables controlled, so cause and effect can be established. Independent measures is needed because the drugs cannot be taken one after the other without carry-over effects, and random allocation spreads participant variables such as severity evenly. The double-blind procedure removes demand characteristics and experimenter bias, improving validity. Measuring at several points shows how quickly each drug works, and the follow-up shows whether improvement lasts. Using both a self-report and a clinician rating gives triangulation. Ethical guidelines are vital because depressed patients are vulnerable.

Psychology and consumer behaviour

6(a) I would carry out a quasi-experiment in primary schools and nurseries.

IV: age, with three groups: 3 to 4 years, 5 to 6 years and 7 to 8 years. Age is naturally occurring, so it is a quasi-experiment with an independent measures design.

Sample: 30 children per age group (90 in total), with roughly equal numbers of boys and girls, chosen by opportunity sampling from schools that agree to take part.

Procedure: each child is tested individually in a quiet room by the same researcher. The child is given 20 cards, each showing a brand logo with the brand name removed: 10 from children's products (such as fast food, cereals and toys) and 10 from adult products (such as cars and banks). The child places each card on a

Question number	
	board showing pictures of 20 products, matching the logo to its product. DV: the number of logos correctly matched, out of 20 (and out of 10 for each type of product). Controls: the same cards, board, instructions and researcher for every child; the order of cards is randomised for each child to avoid order effects; testing at the same time of day. Ethics: informed consent from parents and the head teacher, the child's own agreement, the right to withdraw (a child can stop at any time), no products harmful to children, and the task is presented as a game so it is not stressful. Analysis: mean scores for each age group shown in a bar chart and compared with a Kruskal-Wallis test.
6(b)	Psychologically, brand recognition is the extent to which a consumer can correctly identify a product just by seeing its logo, tag line, packaging or advertising. Fischer et al. (1991) studied brand recognition in children aged 3 to 6 by asking them to match logo cards to products on a game board, and found that recognition increased with age and that many 6-year-olds could match Old Joe the Camel to cigarettes. My study uses the same matching method but extends the age range to see whether recognition keeps increasing, and compares children's and adult products to see whether children recognise brands aimed at them earlier. Methodologically, a quasi-experiment is used because age cannot be manipulated, and independent measures is needed because each child is only one age. A matching task suits young children who cannot read or write, and removing brand names means the DV measures logo recognition, not reading, which improves validity. Testing individually with standardised cards and instructions controls situational variables and makes the study replicable and reliable, and randomising card order prevents order effects. The score out of 20 gives quantitative data for statistical comparison. Ethical guidelines are especially important because the

Question number

participants are young children, so parental consent and the right to withdraw must be respected.

Psychology and health

7(a) I would use a questionnaire completed by adults in family doctors' waiting rooms.

Sample: 200 adults aged 18 to 65, 100 men and 100 women, from five practices in different areas, chosen by opportunity sampling until each gender quota is filled.

The IV is gender.

Type and setting: a paper questionnaire completed alone and anonymously in the waiting room, then posted into a sealed box.

Questions: Section 1 has closed Likert items rated 1 (strongly disagree) to 5 (strongly agree), such as "I often worry that I have a serious illness", "When I notice a pain or change in my body I think it is a sign of disease" and "Even when my doctor tells me I am healthy, I still worry". Some are reverse worded, such as "I rarely think about my health", to reduce response bias. Section 2 asks "How many times have you visited a doctor in the last 12 months?". Section 3 has an open question: "Describe what you do when you notice a new symptom."

Scoring: Likert scores are added to give a hypochondriasis score; higher means more health anxiety.

Analysis: mean scores for men and women compared with a Mann-Whitney U test and shown in a bar chart; open answers analysed for themes.

Reliability: test-retest with a sub-sample after two weeks.

Validity: a pilot study checks clarity, and with consent self-reported visits can be checked against records.

Ethics: informed consent, anonymity, the right to withdraw, and a debrief sheet with contact details for support, as the questions may cause worry.

Question number	
7(b)	Psychologically, the study is based on Barlow and Durand's (1995) description of hypochondriasis as a severe fear of having a serious medical illness, in which minor symptoms or normal bodily sensations are perceived as signs of something more serious, and the anxiety continues even after medical reassurance. A trigger, such as watching a television programme about an illness, can lead to a perceived threat. It often occurs with other disorders such as OCD. My Likert items are written to measure these features: worry about illness, misinterpreting symptoms and not being reassured. Hypochondriasis is a form of misusing health services, so the number of doctor visits is also measured. Methodologically, a questionnaire allows a large sample to be tested quickly and cheaply, and recruiting equal numbers of men and women allows a fair comparison. Closed Likert questions give quantitative data so the genders can be compared statistically, and the open question adds qualitative detail, improving validity. Anonymity reduces social desirability, which matters because men may be less willing to admit health worries. Reverse-worded items reduce response bias, test-retest checks reliability and a pilot study improves clarity. Recruiting in waiting rooms reaches people using health services, the group where hypochondriasis shows. Ethics are addressed with consent and a support debrief.
	Psychology and organisations
8(a)	I would use the factory's attendance records (secondary data) together with return-to-work interviews. Sample: all 300 production workers in one food production factory, with permission from management and the workers' consent. Procedure: absence records are collected for 12 months, recording the day of the week of each absence. Public holidays are removed. When a worker returns, a trained interviewer from the human resources department asks structured

Question number

questions, such as "What was the reason for your absence?", and the reason is coded into one of Blau and Boal's four categories of absence: medical, career-enhancing, normative or calculative.

Variables: the day of the week (Monday to Friday, plus weekend shifts) and the DV, the number of absences on each day, calculated as a percentage of the workers scheduled on that day so that days with more shifts are compared fairly.

Controls: the same definition of an absence (missing a full shift) and the same recording system for the whole year; a full year avoids seasonal effects such as winter illness.

Reliability: two coders independently classify the reasons for absence to check inter-rater reliability.

Ethics: data is anonymised, workers give consent, and records are never used for discipline.

Analysis: the percentage of absences on each day shown in a bar chart and tested with a chi-squared test against the expected equal spread; the type of absence is also compared across days, for example whether Monday absences are mainly calculative.

8(b) Psychologically, the study is based on Blau and Boal (1987), who explained absenteeism through job involvement and organisational commitment and identified four types of absence: medical (genuine illness), career-enhancing (absence to pursue career activities, such as attending a course), normative (a habitual response, where absence is accepted as normal) and calculative (where workers calculate and use the maximum absence permitted). If absences cluster on Mondays and Fridays, this suggests calculative or normative absence to extend the weekend rather than medical absence, so coding the reasons tests this.

Methodologically, company records are objective, quantitative and already collected, so they are cheap, avoid demand characteristics and cover a whole year. Converting

Question number

absences to percentages controls for different numbers of shifts on each day, improving validity. Return-to-work interviews add the reason, giving qualitative data that explains the pattern, and structured questions make them standardised and reliable. Inter-rater checks make the coding consistent. Studying one factory gives high ecological validity, although it limits generalisation to other organisations. Anonymity protects workers, which also makes honest answers about absence more likely.

Section C

Psychology and abnormality

9 The nature versus nurture debate asks whether behaviour is caused by biological factors we are born with, such as genes and neurochemicals (nature), or by experience and learning (nurture).

There is evidence that the debate is relevant to depression. Nature explanations include genetic and neurochemical research: Oruc et al. (1997) found that patients with bipolar disorder often had first-degree relatives with affective disorders, and that serotonin receptor genes were linked to risk in females. The neurochemical explanation says low serotonin causes depression. Nurture explanations include Beck's cognitive triad, where negative thoughts about the self, the world and the future are learned, and learned helplessness (Seligman), where repeated uncontrollable bad events teach a person that they cannot change their situation. Because the debate is reductionist, each side can be studied precisely. Most importantly, it determines the type of treatment: a nature cause suggests drugs such as SSRIs or MAOIs, while a nurture cause suggests cognitive restructuring or REBT (Ellis), which changes irrational beliefs.

However, there are reasons to agree that the debate is irrelevant. People suffer

Question number

from depression whether the cause is nature or nurture, and their symptoms will be treated either way. The debate is theoretical: in practice, drugs and CBT are often combined. For some people the cause may be mainly nature, and for others mainly nurture, so asking which one is "the" cause is too simple. The diathesis-stress approach suggests both interact: a genetic vulnerability may only lead to depression when stressful life events occur. Also, nature and nurture are hard to separate, since negative thinking may change brain chemistry and vice versa.

In conclusion, I disagree that the debate is irrelevant, because it has guided research and the choice of treatment. However, framing it as nature or nurture is unhelpful; an interactionist view, where both contribute, best explains depression and leads to the most effective combined treatment.

Psychology and consumer behaviour

10 Advertising media include television, radio, newspapers and magazines, posters, the internet and product placement in films.

There is some support for the view that television is the most effective medium. Large numbers of most populations watch television, so it reaches a mass audience. It can be general or focused on one region or channel, and it combines moving pictures, sound and words, so it has more impact than a poster. It can use credible or attractive sources such as celebrities, which the Yale model of communication (Hovland et al., 1953) says increases persuasion. Cowpe (1989) showed television adverts reduced chip pan fires in the regions where they were shown, showing its power to change behaviour.

However, there is strong evidence against the idea that there is only one effective way. Some people do not have a television, watch streaming services or skip adverts. Other media can be more effective for particular audiences. Fischer et al.

Question number

(1991) found many 6-year-olds recognised Old Joe the Camel, a character used in print and poster advertising, so non-television media can be highly effective even with children. Auty and Lewis (2004) found that children who saw a film clip with Pepsi placed in it were more likely to choose Pepsi, showing product placement works. Snyder and DeBono (1985) found that magazine-style adverts work differently for different people: high self-monitors preferred image-based adverts, while low self-monitors preferred adverts about product quality. This shows the best method depends on the audience, not the medium alone. Specific media, such as a magazine or a poster on a student campus, can also target a particular group more cheaply than television.

In conclusion, I disagree with the statement. Television is very powerful for mass audiences, but research shows that print, posters and product placement can also be effective, and that the best method depends on the product and the audience, so a combination of media is the most effective strategy.

Psychology and health

11 Psychological measures of stress are self-reports, such as the Social Readjustment Rating Scale (Holmes and Rahe, 1967) or the Hassles and Uplifts Scale (Kanner et al., 1981). Physiological measures include recording devices, such as heart rate and galvanic skin response (GSR) monitors used by Geer and Maisel (1972), and sample tests, such as the salivary cortisol measured by Evans and Wener (2007) in train commuters.

There is some support for the statement. Psychological measures allow people to explain how they feel and why, and stress is a subjective experience, so this gives them construct validity. For example, the Hassles scale shows that small daily problems matter, which a hormone level cannot reveal. Physiological measures are correlational, and a raised heart rate could be caused by exercise, caffeine or

Question number

excitement rather than stress, so without an explanation the number means little. Using one physiological measure alone is also reductionist.

However, there is strong evidence against the statement. Physiological data is objective and not open to bias, such as social desirability or poor memory, which affect self-reports: the SRRS asks people to recall events over a year, which may be inaccurate. Recording devices give consistent, reliable measurements on a standard scale, for example a sphygmomanometer measures blood pressure the same way every time. Human physiology is similar across cultures, so physiological measures generalise better, whereas the SRRS life events and their weightings are culturally biased. Geer and Maisel's use of GSR and heart rate gave precise, quantitative data showing that perceived control reduced stress.

In conclusion, I disagree with the statement. Physiological measures are generally more reliable and objective, while psychological measures capture the subjective meaning of stress, which can make them more valid in some ways. Neither is better overall, which is why the best studies, such as Evans and Wener (2007), combine both cortisol and self-reported stress.

Psychology and organisations

12 Belbin identified nine team roles in three groups: thought-related (plant, monitor evaluator, specialist), people-related (coordinator, teamworker, resource investigator) and action-related (shaper, implementer, completer finisher). People are categorised using the Belbin Team Inventory, which has a self-perception questionnaire and observer assessments.

There is some support for the statement. The inventory has features of a psychometric test, such as standardised scoring, and it developed from a number of psychometric tests, so it measures people in a consistent way. Classifying people into types is exactly what it does, and the nine roles cover thinking, social

Question number

and practical contributions, so managers can build balanced teams and avoid having, for example, too many plants and no completer finisher. It is widely used in organisations, which suggests it is useful in practice.

However, there is strong evidence that it is not perfect. Belbin himself says it measures "a tendency to behave, contribute and interrelate with others in a particular way", not a fixed type, so people may play different roles in different teams. The self-perception test is only as good as a person's sense of personal realism: some answer in terms of how they would like to contribute rather than how they really behave, which is why observer assessments are added. Furnham et al. (1993) assessed the self-perception inventory and found weak psychometric properties, such as low reliability. Categorising people is also reductionist, as roles overlap and people have mixed strengths. Finally, the inventory analyses how people work together, while most tests focus on individuals, which makes it hard to validate.

In conclusion, I disagree that Belbin's roles are the perfect way to categorise people. They are a useful, practical framework for building balanced teams, but because of self-report bias, questionable reliability and the fact that roles change with the situation, they should be used flexibly and combined with observation.

Cambridge Assessment International Education

Cambridge International AS & A Level

PSYCHOLOGY **9990/43**

Paper 4 Specialist Options: Application **May/June 2022**

1 hour 30 minutes

O8131862O

You must answer on the enclosed answer booklet.

You will need: Answer booklet (enclosed)

INSTRUCTIONS

- Answer **four** questions in total:
  - Answer questions from **two** options.
    - Section A: answer **two** questions.
    - Section B: answer **one** question.
    - Section C: answer **one** question.
- Follow the instructions on the front cover of the answer booklet. If you need additional answer paper, ask the invigilator for a continuation booklet.

INFORMATION

- The total mark for this paper is 60.
- The number of marks for each question or part question is shown in brackets [ ].

This document has **8** pages. Any blank pages are indicated.

© UCLES 2022

[Turn over

stemup.co.uk

Section A

You must answer **two** questions from this section.
Answer all the parts for the two questions you choose.

Psychology and abnormality

1 In a case study by Glover, a 56-year-old married woman who had been shoplifting every day for 14 years attended treatment sessions. Her treatment used covert sensitisation with aversive imagery.

(a) Explain what is meant by the term 'aversive imagery'. [2]

(b) Suggest **two** strengths of the use of a face-to-face interview by a therapist treating a patient with kleptomania. [4]

(c) Suggest why biochemical treatments might **not** be effective when treating kleptomania. [4]

(d) Discuss whether generalisations can or cannot be made from this case study. You should consider both sides of the argument and include a conclusion. [5]

Psychology and consumer behaviour

2 Dayan and Bar-Hillel (2011) conducted research on how the position of a food item on a menu influenced what participants ordered to eat. They compared four versions of menu item order. They described their sample and sampling technique in Study 1 as:

- 240 participants
- from the Hebrew University, Jerusalem
- all were students
- aged 19–35 years
- 52% female, 48% male
- recruited individually from around the campus.

(a) Suggest which sampling technique could have been used in Study 1. [2]

(b) **(i)** Suggest **one** strength of the sample used in this study. [2]

(ii) Suggest **one** weakness of the sample used in this study. [2]

(c) **(i)** Identify **two** of the versions of menu item order used in Study 1. [2]

(ii) Explain why participants were randomly allocated to a version of menu item order. [2]

(d) Discuss the strengths and weaknesses of using a laboratory experiment to study menu item order. You should include a conclusion in your answer. [5]

© UCLES 2022

stemup.co.uk

Psychology and health

3 McKinstry and Wang (1991) investigated the way in which doctors' clothing influenced patients' confidence in doctors. Patients were shown photographs of different styles of doctors' clothing. Data was gathered by an interviewer asking closed questions in a structured interview.

(a) Explain what is meant by a 'structured interview'. [2]

(b) Give **one** strength and **one** weakness of asking closed questions as used in this study. [4]

(c) **(i)** Suggest how the quantitative data gathered for the styles in relation to confidence could be analysed. [2]

(ii) Suggest how the quantitative data gathered for the styles in relation to confidence could be shown on a graph. [2]

(d) Discuss the strengths and weaknesses of using photographs to gather information about a doctor's clothing. You should include a conclusion in your answer. [5]

Psychology and organisations

4 In just one year in the UK there were 21 deaths and 1372 injuries caused by chip pan fires. Cowpe (1989) reported on a safety campaign where two television advertisements were shown in ten regions of the UK between 1976 and 1982.

(a) Suggest the type of experiment used in this study. [2]

(b) Explain how **two** safety promotion strategies were used in the television advertisements for this campaign. [4]

(c) **(i)** Explain how the effectiveness of the campaign was measured. [2]

(ii) Suggest **one** strength of gathering data in this way. [2]

(d) Discuss the strengths and weaknesses of using television to promote safety in organisations. You should include a conclusion in your answer. [5]

© UCLES 2022

stemup.co.uk

Section B

You must answer **one** question from this section.

Psychology and abnormality

5 **(a)** Design a study to investigate the effectiveness of biochemical treatments for impulse control disorder. [10]

(b) Explain the psychological and methodological evidence on which your study is based. [8]

Psychology and consumer behaviour

6 It is suggested that people have cognitive maps of where individual food items are located in retail environments, such as supermarkets.

(a) Design a study to investigate whether people use cognitive maps to locate individual food items in a supermarket. [10]

(b) Explain the psychological and methodological evidence on which your study is based. [8]

Psychology and health

7 **(a)** Design a study to investigate whether doctor-centred or patient-centred practitioners are more likely to make incorrect diagnoses. [10]

(b) Explain the psychological and methodological evidence on which your study is based. [8]

Psychology and organisations

8 **(a)** Design a study using a questionnaire to investigate which non-monetary reward is more effective for workers in an organisation. [10]

(b) Explain the psychological and methodological evidence on which your study is based. [8]

© UCLES 2022

stemup.co.uk

Section C

You must answer **one** question from this section.

Psychology and abnormality

9 *'Psychometric measures used to assess obsessive-compulsive disorder (OCD) provide therapists with no useful information.'*

To what extent do you agree with this statement? Use examples of research you have studied to support your answer. [12]

Psychology and consumer behaviour

10 *'Studies using an fMRI, such as those on pre-cognitive decisions, are of no value in understanding consumer behaviour.'*

To what extent do you agree with this statement? Use examples of research you have studied to support your answer. [12]

Psychology and health

11 *'A biochemical test is the only accurate way to know that a person has adhered to a request to take their medication.'*

To what extent do you agree with this statement? Use examples of research you have studied to support your answer. [12]

Psychology and organisations

12 *'There is no difference between Alderfer's ERG theory and Maslow's hierarchy of needs.'*

To what extent do you agree with this statement? Use examples of research you have studied to support your answer. [12]

© UCLES 2022

stemup.co.uk

BLANK PAGE

© UCLES 2022

stemup.co.uk

BLANK PAGE

© UCLES 2022

stemup.co.uk

BLANK PAGE

Permission to reproduce items where third-party owned material protected by copyright is included has been sought and cleared where possible. Every reasonable effort has been made by the publisher (UCLES) to trace copyright holders, but if any items requiring clearance have unwittingly been included, the publisher will be pleased to make amends at the earliest possible opportunity.

To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced online in the Cambridge Assessment International Education Copyright Acknowledgements Booklet. This is produced for each series of examinations and is freely available to download at www.cambridgeinternational.org after the live examination series.

Cambridge Assessment International Education is part of Cambridge Assessment. Cambridge Assessment is the brand name of the University of Cambridge Local Examinations Syndicate (UCLES), which is a department of the University of Cambridge.

© UCLES 2022

stemup.co.uk

Surname	StemUp Model Solution
Answer booklet	Write the question number in the left-hand margin at the start of each answer.

Question number

Answers for paper 9990/43

Section A

Psychology and abnormality

1(a) Aversive imagery is when the patient imagines a very unpleasant (aversive) scene and pairs it with the unwanted behaviour, so the behaviour becomes associated with negative feelings and is less likely to be repeated (a form of punishment). In Glover's case study the woman imagined feeling sick and vomiting as she approached an item in a supermarket, with other shoppers watching her, until she put the item back and left the store.

1(b) One strength is that the therapist can observe the patient's non-verbal behaviour, such as facial expression, hesitation or fidgeting, while she answers questions. This helps diagnosis and therapy, for example showing whether she is being honest about how often she shoplifts or which situations make the urge to steal strongest.

A second strength is that a face-to-face interview is more personal than a telephone interview. The therapist gives time to the patient and can build rapport and trust, so a patient with kleptomania, who may feel ashamed of stealing, is more likely to disclose details and to ask questions about her disorder and treatment.

1(c) There is no drug that specifically stops a person from stealing. Drugs can only

Question number

reduce general symptoms such as tension or anxiety; they do not remove the cause. Kleptomania involves a learned cycle of rising tension before the theft and pleasure or relief afterwards, so if this learned behaviour is not changed the patient is likely to relapse when the drug is stopped.

Also, the patient may not adhere to the medication, for example if she does not see stealing as a medical problem, and long-term use of a drug may cause side effects or dependence, creating extra problems.

1(d) Generalisations can be made because covert sensitisation is based on conditioning principles that apply to everyone, so the technique of pairing the urge to steal with an aversive image could be generalised to other patients with kleptomania. The same technique could also be generalised to other impulse control disorders, such as pyromania or pathological gambling, by changing the imagined scene.

However, generalisations cannot be made about the specific details. This was one 56-year-old married woman who had stolen daily for 14 years, so her history, motivation and response to treatment are unique to her. Also, the particular aversive images of nausea and vomiting will not be unpleasant enough for everyone; other patients would need different images personal to them, so the exact procedure cannot be generalised.

In conclusion, the general technique can be generalised but the individual procedure cannot, so the case study shows covert sensitisation is worth trying, but it should be tested on a larger sample with images tailored to each patient before it is assumed to work for everyone.

Psychology and consumer behaviour

2(a) An opportunity sample, because the participants were students recruited individually from around the campus, so they were simply the people who were

Question number	
	available and willing at the time the researchers approached them.
2(b)(i)	The sample had a nearly equal gender balance (52% female, 48% male), so menu choices were not biased towards one gender and the findings about item position can be generalised to both men and women.
2(b)(ii)	All participants were students from one university in Jerusalem aged only 19 to 35, so the findings may not generalise to older diners or people from other cultures, who may read and choose from menus differently.
2(c)(i)	The Baseline menu, which had the items in an arbitrary order, and the Mirror menu, which completely reversed the Baseline order within each food category.
2(c)(ii)	Random allocation gives each participant an equal chance of being in any of the four menu conditions, so participant variables such as food preferences or hunger are spread evenly. This avoids systematic bias, so any difference in what was ordered is due to the position of the items.
2(d)	A strength is control. The IV (the order of items on the menu) is manipulated and the DV (the item chosen) is measured while extraneous variables, such as the items, descriptions and prices, are kept the same, so cause and effect can be established. A second strength is that one variable, position on the menu, can be isolated and studied on its own, and the study is easy to replicate to check reliability. A weakness is low ecological validity. Real ordering happens in restaurants with waiters, friends, smells, specials and real money, so choices on a menu in a study may not reflect what people really order. Also, participants know they are in a study, so they may respond to demand characteristics, for example avoiding the

Question number	
	first item if they guess the aim, which lowers validity. In conclusion, a laboratory experiment is the best way to show that menu position alone can change choices, but because real dining involves many other variables, its findings should be checked with a field experiment in a real restaurant before being applied by menu designers.
	Psychology and health
3(a)	A structured interview is one where the interviewer asks every participant the same pre-set questions in the same order, with no extra questions added. For example, McKinstry and Wang asked every patient which doctor they would feel happiest about seeing for the first time.
3(b)	A strength is that closed questions produce quantitative data, because every patient answers on the same rating scale. This means the confidence scores for the different styles of clothing, such as a formal suit or casual clothes, can be compared statistically, and the study can easily be replicated. A weakness is that patients cannot explain the reason for their answer. A patient might prefer the doctor in formal clothes because of the doctor's age or expression rather than the clothing, but closed questions give no opportunity to add this information, so validity may be lower.
3(c)(i)	Calculate the mean confidence score for each style of clothing (for example white coat, formal suit, casual), then compare the means to see which style gives the most confidence, and use the range to show how varied the patients' ratings were.
3(c)(ii)	A bar chart, with a separate bar for each style of doctor's clothing on the x-axis and the mean confidence score on the y-axis, so the style with the tallest bar

Question number

gives patients the most confidence.

3(d) A strength is standardisation: in a photograph the doctor's expression, pose and background stay exactly the same, so the only thing that changes between photographs is the clothing, which increases validity. A second strength is that the study can be replicated with the same photographs, and they can be shown to patients in many different practices, not just one place.

A weakness is that photographs are static, while real doctors move, speak and show a manner and tone of voice. Patients' confidence in a real consultation depends on these too, so judging clothing from a photograph has low ecological validity. Also, the people shown in the photographs must give consent for their images to be used, and patients may focus on the person's face or age rather than the clothing.

In conclusion, photographs are the most controlled way to isolate the effect of clothing, so they are useful for a first study, but the findings should be checked with real doctors in real consultations before practices make rules about what doctors wear.

Psychology and organisations

4(a) A field experiment, because the IV (whether the television advertisements were shown in a region) was manipulated by the researchers, and it took place in a real-life setting, people's homes in ten regions of the UK, with the DV being the number of chip pan fires.

4(b) One strategy was prevention: the advertisements showed people how to avoid a chip pan fire starting, for example by not overfilling the pan with oil and never leaving it unattended while it is heating.

Question number	
	A second strategy was containment: the advertisements showed the correct procedure if a fire does start, such as turning off the heat, covering the pan with a damp tea towel and leaving it to cool, and showed that throwing water on it or carrying the pan outside is wrong. Showing a real fire and people's shocked reactions also used fear arousal to make the message memorable.
4(c)(i)	By the number of chip pan fires recorded in official fire brigade statistics in each region, comparing the number of fires before the advertisements with the number during and in the months after they were shown.
4(c)(ii)	The data is objective: a fire brigade call-out either happened or it did not, so it cannot be biased by social desirability, unlike asking people whether they had a chip pan fire. It is also quantitative, so the number of fires in different regions can be compared.
4(d)	A strength is reach: large numbers of the population, including most workers, watch television, so a safety message reaches far more people than a poster in one workplace. A second strength is impact: television can show a real hazard, such as a fire, and model the correct behaviour, so it can arouse fear and show people exactly what to do, which makes the message memorable. It can also be targeted at one region or channel. A weakness is that some people do not have a television, watch streaming services without adverts or skip the adverts, so they never see the message. Also, television messages are general, so they cannot cover the specific hazards of one organisation, such as a particular machine, and the organisation cannot check whether workers understood or followed the advice. In conclusion, television is effective for raising general safety awareness, as the fall in chip pan fires suggests, but within an organisation it should be backed by

Question number

on-site methods such as training or a token economy for safe behaviour (Fox et al., 1987), which target its own hazards.

Section B

Psychology and abnormality

5(a) I would carry out a randomised, double-blind, placebo-controlled experiment on pathological gambling, an impulse control disorder, testing the drug nalmefene, an opioid antagonist.

Sample: 80 adults aged 18 to 65 diagnosed with pathological gambling by a psychiatrist using DSM criteria, recruited as a volunteer sample through advertisements in gambling-support clinics. People with another disorder, such as bipolar disorder, or already taking medication would be excluded.

Design and IV: an independent measures design. Participants would be randomly allocated to two conditions: nalmefene tablets or identical-looking placebo tablets, taken daily for 16 weeks. Neither the participants nor the psychiatrists rating them would know who had which tablet (double blind).

DV: the urge to gamble, measured with the PG-YBOCS, a version of the Yale-Brown scale for gambling, rated by a psychiatrist at the start and every four weeks. Participants would also keep a weekly diary of the number of gambling episodes and the money spent, giving a second, quantitative measure.

Controls: same dose, same timing, same rating scale and same interview procedure for everyone; no other treatment during the study.

Reliability and validity: two psychiatrists would rate each patient independently to check inter-rater reliability. Validity is improved by the placebo group and double blind, which control expectations and experimenter bias.

Ethics: fully informed consent, the right to withdraw at any time, monitoring of

Question number	
	side effects, confidential data, and a full debrief, with the placebo group offered the real treatment at the end.
	Analysis: the mean reduction in PG-YBOCS score would be compared for the two groups using a Mann-Whitney U test, and shown in a bar chart.
5(b)	Psychologically, the study is based on the biochemical explanation of impulse control disorders. Gambling gives a "high" through opioids and dopamine in the brain's reward system. Nalmefene is an opioid antagonist which blocks opioid receptors, so it should reduce the pleasure of gambling and therefore the urge to gamble. Grant et al. (2008) tested nalmefene on pathological gamblers: participants were assessed to make sure they were suitable, then randomly allocated to nalmefene or a placebo, and the urge to gamble was measured with the PG-YBOCS. My study copies this design so the results can be compared with theirs.
	Methodologically, a laboratory-style clinical experiment allows the IV (drug or placebo) to be manipulated and other variables controlled, so a cause and effect conclusion about the drug can be drawn. Independent measures is needed because a participant cannot take the drug and not take it, and random allocation spreads participant variables, such as severity, evenly. The placebo controls for the expectation of getting better, and the double-blind procedure stops participants and raters showing demand characteristics or bias. Using a standardised clinician-rated scale gives quantitative, reliable data, while the gambling diary adds a measure of real behaviour, improving validity. Ethical guidelines are followed because the participants are vulnerable patients who must be protected from harm.
	Psychology and consumer behaviour
6(a)	I would carry out a quasi-experiment in a large supermarket, comparing regular

Question number

shoppers with people who have never shopped there.

Sample: 30 regular shoppers (who shop there at least once a week) and 30 new shoppers, aged 18 to 65, selected by opportunity sampling at the entrance, with permission from the store manager.

IV: familiarity with the store (regular or new). This is naturally occurring, so the design is independent measures.

Procedure and DVs: before entering, each participant is given a blank outline of the store and asked to draw a sketch map showing where 10 named items are, such as milk, rice, bread and toothpaste. Score: the number of items placed in the correct aisle (out of 10). Each participant is then given the same list and asked to collect the items. A researcher follows them (a non-participant, overt observation) and records the time taken and the number of wrong aisles entered, marking their route on a store plan.

Controls: the same 10 items for everyone, testing on weekday mornings when the store is quiet, and no changes to the store layout during the study.

Reliability: two researchers would score each sketch map independently to check inter-rater reliability.

Ethics: informed consent, the right to withdraw, data kept anonymous, and participants keep any items they want to buy as normal.

Analysis: mean sketch-map scores and mean search times for each group compared with a Mann-Whitney U test; if regular shoppers draw more accurate maps and find items faster, this suggests they use a cognitive map.

6(b) Psychologically, a cognitive map is a mental representation of a spatial layout that a person uses to find their way around. MacKay and Olshavsky (1975) studied consumers' cognitive maps of retail locations, and one way they measured them was by asking people to draw sketch maps, which is why my study uses a sketch-map task. If people have a cognitive map of a supermarket, it should be

Question number

more accurate the more often they shop there, so familiarity is the IV. Gil et al. (2009) used CCTV to record shopper movement patterns, showing that routes through a store can be recorded, which supports observing and plotting each participant's route.

Methodologically, a quasi-experiment is used because familiarity cannot be manipulated; it already exists. Independent measures is needed as a person cannot be both a regular and a new shopper. Using two DVs, a sketch map (what people know) and search time (what people do), is a form of triangulation, so the conclusion is more valid than with one measure. Carrying out the study in a real supermarket gives high ecological validity, while the controls (same items, same time of day) reduce extraneous variables. Quantitative scores allow statistical comparison, and checking inter-rater reliability makes the scoring of maps consistent. Opportunity sampling is practical at a store entrance, and consent protects participants.

Psychology and health

7(a) I would use a field study with simulated patients (trained actors) visiting family doctors, combined with a structured observation of the consultations.

Sample: 40 family doctors (GPs) from different practices, recruited as volunteers. They agree that over the next six months some of their patients may be actors, but do not know which.

Step 1, classifying doctors: with consent, 10 of each doctor's ordinary consultations are audio-recorded. Two trained observers code them using behavioural categories, such as doctor-centred (closed questions, focusing on the first problem mentioned, giving instructions) and patient-centred (open questions, listening, asking the patient's views). Each doctor is classed as doctor-centred or patient-centred by their main style. This gives the IV, so it is a quasi-experiment with independent

Question number

groups.

Step 2, measuring errors: each doctor is visited by four actors with scripted symptoms: two have a known condition (such as the symptoms of a migraine) and two are healthy but describe a minor worry. DV: the number of incorrect diagnoses out of four, counted as false positives (diagnosing an illness in a healthy patient) and false negatives (missing the real condition).

Controls: all actors use the same scripts, the same symptoms and the same appointment length.

Reliability: two observers code the recordings, checking inter-rater reliability.

Ethics: consent from doctors, patients consent to being recorded, confidentiality of doctors' results, and a debrief explaining which visits were actors.

Analysis: percentage of incorrect diagnoses for each style compared with a chi-squared test.

7(b) Psychologically, the study is based on Byrne and Long (1976), who analysed over 2000 tape-recorded GP consultations and identified a range of styles from doctor-centred, where the doctor controls the consultation, to patient-centred, where the patient is involved. My coding categories are based on their behaviours. Savage and Armstrong (1990) also compared a directive (doctor-centred) and a sharing (patient-centred) style. A doctor-centred doctor who asks only closed questions may miss information and so make more diagnostic errors. The study measures both types of error: Rosenhan (1973) showed that doctors are more inclined to call a healthy person sick (a false positive, which he called a type 2 error) than to call a sick person healthy, which is why healthy actors are included. Methodologically, a structured observation with clear behavioural categories gives quantitative data and can be checked for inter-rater reliability. Using simulated patients means the correct diagnosis is known, so errors can be measured objectively, and the same scripts act as a control, making the visits standardised.

Question number

Real consultations give high ecological validity, and doctors not knowing which patients are actors reduces demand characteristics. A quasi-experiment is used because style cannot be manipulated. Ethics are addressed with prior consent and a full debrief.

Psychology and organisations

8(a) I would use a questionnaire given to employees of a large supermarket chain.

Sample: 200 employees chosen by stratified sampling, so the proportions of shop-floor staff, supervisors and managers, and of men and women, match the organisation.

Type and setting: an anonymous online questionnaire completed during work time in a private room, taking about 15 minutes.

Questions: Section 1 has closed Likert-scale items for five non-monetary rewards: praise, respect, recognition, empowerment and a sense of belonging, for example "Being praised by my manager makes me work harder" and "Being given the power to make decisions in my job motivates me", rated from 1 (strongly disagree) to 5 (strongly agree). Some items are reverse worded, such as "Being thanked in public makes no difference to my effort", to reduce response bias. Section 2 asks workers to rank the five rewards from 1 (most motivating) to 5. Section 3 has an open question: "Describe a time when a reward that was not money made you work harder."

Scoring and analysis: the mean Likert score and the median rank are calculated for each reward and shown in a bar chart; the reward with the highest mean is the most effective. Open answers are analysed for themes.

Reliability: the questionnaire is given again two weeks later to check test-retest reliability.

Validity: a pilot study checks the questions are clear.

Question number

Ethics: informed consent, anonymity, the right to withdraw, and managers never see individual answers.

8(b) Psychologically, the study is based on the types of non-monetary reward: praise (being told you have done well), respect (being valued as a person), recognition (achievement being publicly acknowledged, such as employee of the month), empowerment (being given responsibility and decision-making power) and a sense of belonging (feeling part of the team). These link to need theories: Maslow (1970) argues that people are motivated by belonging and esteem needs, which belonging, praise and recognition meet, and Herzberg's two-factor theory treats recognition and responsibility as motivators. The study tests which of these rewards is actually most effective.

Methodologically, a questionnaire allows a large sample to be asked quickly, and stratified sampling makes it representative of the whole organisation. Closed Likert questions give quantitative data so the rewards can be compared with means, while ranking forces a clear choice. The open question gives qualitative data explaining why a reward works, adding validity. Anonymity reduces social desirability bias, which is likely when workers are asked about their managers, and reverse-worded items reduce response bias. Test-retest checks reliability, and a pilot study improves validity by removing unclear questions. Ethical guidelines protect workers from any consequences at work.

Section C

Psychology and abnormality

9 Psychometric measures are standardised tests used to assess a disorder. Two used for OCD are the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) and the

Question number

Maudsley Obsessive-Compulsive Inventory (MOCI). The Y-BOCS is clinician-rated: the therapist asks about obsessions and compulsions separately, rating items such as time spent, distress and resistance from 0 to 4, giving a total out of 40. The MOCI is a self-report of 30 true or false statements with subscales such as checking, washing, slowness and doubting.

There is strong evidence against the statement. These measures give the severity of the disorder as a number, so a therapist can decide whether treatment is needed and which treatment suits the patient. They also show the specific features of a person's OCD: the MOCI subscales show whether the problem is mainly washing or checking, so exposure and response prevention can target the right behaviour. Because the data is quantitative, scores can be compared with other people with the disorder and measured again during therapy to track progress. For example, in Lehmkuhl et al.'s (2008) case study of Jason, a 12-year-old with autism and OCD, a children's version of the Y-BOCS was used to show that his symptoms fell after CBT.

However, there are limitations that partly support the statement. Self-reports such as the MOCI rely on honesty, and people with OCD may hide embarrassing obsessions because of social desirability. Fixed answers such as true or false, or a rating scale with a neutral option, may not represent what the person really thinks. The measures are nomothetic: they assume people with OCD share similar symptoms, so an unusual obsession may be missed. A clinical interview can reveal more, such as the meaning of the obsession or what triggered it, which gives qualitative detail a score cannot.

In conclusion, I largely disagree with the statement. Psychometric measures clearly provide useful information about severity, type of symptom and progress, which is why therapists use them. Their weaknesses mean they should not be used alone, but combined with a clinical interview they give therapists a valid and reliable picture of the patient.

Question number

Psychology and consumer behaviour

10 fMRI measures brain activity by detecting changes in blood flow while a person does a task. Knutson et al. (2007) used fMRI to study pre-cognitive decisions about buying. Participants in the scanner were shown a product, then its price, and then chose whether to buy it. Activation in the nucleus accumbens when the product was shown predicted that the person would buy it, while activation in the insula when the price was seen as too high predicted that they would not buy. Activity in the medial prefrontal cortex was linked to the price. This suggests the decision starts in the brain before the person is aware of it.

There is evidence that such studies are of value. fMRI is scientific equipment producing objective, quantitative data that is not affected by social desirability, unlike asking consumers why they bought something. Studying how the mind makes decisions is always useful, and Knutson's findings show that emotion (anticipated pleasure) and the "pain" of paying both affect purchasing. This data could be useful to sellers: if they know which features activate reward areas or what price triggers the insula, they can design and price products people are more likely to buy.

On the other hand, there are reasons to agree with the statement. Lying still in a noisy scanner looking at pictures has low ecological validity compared with shopping in a store with friends, music and time pressure. Samples are small and fMRI is expensive, so findings are hard to generalise. The data is correlational: brain activity is linked to a choice but does not explain why. It is also reductionist, ignoring social and cultural factors. Finally, a decision made before awareness may be irrelevant if the person then acts on it; simply observing what people buy is also factual and is cheaper.

In conclusion, I disagree that fMRI studies have no value, as they give unique

Question number

objective evidence about how buying decisions are made. However, their value is limited by low ecological validity and cost, so they are most useful when combined with real-world methods such as observation of actual purchases.

Psychology and health

11 Adherence means following medical advice, such as taking medication as prescribed. A biochemical test, such as a blood or urine test, detects whether the drug, or a marker added to it, is in the body.

There is support for the statement. Biochemical tests are objective and give quantitative data that the patient cannot manipulate: if the drug is not in the blood, it has not been taken. Roth and Caron (1978) used a biochemical measure of whether patients with ulcers took their antacid medication and found that both patients and doctors overestimated adherence compared with this objective measure. This shows that self-report is not accurate, as patients may forget or give socially desirable answers, such as in self-report questionnaires (for example Riekert and Drotar, 1999).

However, there are strong arguments against the statement. A biochemical test may not be valid: a patient could take the drug only on the day before the test, and the test shows that the drug was taken recently, not that it was taken correctly every day. Tests are invasive (blood samples), costly and time-consuming, so in practice they are rarely done. Also, there are other objective measures. Chung and Naya (2000) used an electronic pill counter (TrackCap), which records each time the bottle is opened, giving a detailed record of when doses were taken, which a single blood test cannot. Repeat prescriptions, as used by Sherman et al. (2000), show whether a patient collects their medication on time, without any invasive procedure. However, these also have limits: opening a bottle or collecting a prescription does not prove the pill was swallowed.

Question number

In conclusion, I disagree that a biochemical test is the only accurate way. It is the most direct measure of whether a drug is in the body, but it only shows one moment in time. The most accurate picture comes from combining methods, for example a biochemical test with electronic pill counting, so each method's weaknesses are covered by the other.

Psychology and organisations

12 Maslow's (1970) hierarchy of needs is a need theory of motivation in which lower needs (physiological, safety) must be met before higher needs (belonging, esteem, self-actualisation), with later versions adding cognitive, aesthetic and transcendence needs to make eight levels. Alderfer's (1972) ERG theory has three needs: existence, relatedness and growth.

There is some support for the statement. Both are need theories, and both say workers are motivated by unmet needs. ERG theory is based on Maslow's needs: existence combines physiological and safety needs, relatedness covers belonging and the social part of esteem, and growth covers self-esteem and self-actualisation. Both recognise the importance of the same basic needs, such as pay and safe conditions, and both suggest the same practical advice for managers, such as meeting basic needs first and then giving opportunities for development.

However, there are important differences, so I mostly disagree. Maslow has eight (or five) levels while Alderfer has three, and combining two needs into one may ignore features of each. Maslow's hierarchy is fixed: a person moves up only when a lower need is satisfied. Alderfer says several needs can motivate at the same time, so a worker may want growth while existence needs are not fully met. Alderfer also added the frustration-regression principle: if a growth need is frustrated, the worker regresses and puts more effort into relatedness, for example socialising more at work, which Maslow does not include. Maslow's later needs,

Question number

such as aesthetic and transcendence needs, do not appear in ERG theory. ERG theory was developed from studies of workers in organisations and is more flexible, so it explains individual differences, for example why people in different cultures rank needs differently, better than Maslow's fixed order.

In conclusion, I disagree with the statement. ERG theory grew out of Maslow's hierarchy and they share the same basic types of need, but ERG theory differs in its number of needs, its flexible order and the frustration-regression principle, making it a different and more applicable theory for the workplace.