



Oxford Cambridge and RSA

A Level Psychology

H567/03 Applied psychology

Thursday 14 June 2018 – Morning

Time allowed: 2 hours



You must have:

- the OCR 12-page Answer Booklet (OCR12 sent with general stationery)

INSTRUCTIONS

- Use black ink.
- Section A: Answer **all** the questions in this section.
- Section B: Choose **two** options and answer **all** parts of the question in the options you have chosen.
- Write your answers in the Answer Booklet. The question number(s) must be clearly shown.
- Do **not** write in the barcodes.

INFORMATION

- The total mark for this paper is **105**.
- The marks for each question are shown in brackets [].
- Quality of extended responses will be assessed in questions marked with an asterisk (*).
- This document consists of **4** pages.

SECTION A: Issues in mental health

Answer **all** the questions.

- 1** **(a)** Outline a non-biological treatment of **one** specific disorder. **[5]**
- (b)** Discuss **one** strength and **one** weakness of the non-biological treatment referred to in your answer to **1(a)**. **[6]**
- 2** **(a)** In the key research by Szasz (2011), the author states that “there is no such thing as mental illness.” How does Szasz defend this claim? **[3]**
- (b)** Evaluate Szasz’s view that “there is no such thing as mental illness.” **[5]**
- 3** Explain how the key research by Rosenhan (1973) contributes to an understanding of individual, social and cultural diversity. **[6]**
- 4*** To what extent are explanations of mental illness determinist? **[10]**

SECTION B: Options

Choose **two** options and answer **all** parts of the question in the options you have chosen.

OPTION 1**Child psychology**

- 5 (a)* Use the key research by Van Leeuwen et al. (2008) to explain the effect of biological factors upon intelligence. [10]

- (b)* Discuss methodological issues involved when researching intelligence. [15]

- (c)* Nasra is manager of a company that designs computer games. She would like to get job applicants doing intelligence tests. The results from these tests would help her decide which candidates to interview.

What advice might a psychologist provide to Nasra's company about the best method for assessing the intelligence of job applicants? [10]

OPTION 2**Criminal psychology**

- 6 (a)* Use the key research by Dixon et al. (2002) to explain how the characteristics of a defendant can affect whether they are found guilty. [10]

- (b)* Discuss ethical considerations of research into psychology and the courtroom. [15]

- (c)* Alan has been put in charge of the court system for his country. He is concerned that in too many cases juries are reaching the wrong verdict. Alan would like to improve jury decision-making.

What might a psychologist suggest about how Alan can improve jury decision-making? [10]

OPTION 3**Environmental psychology**

- 7 (a)* Explain how the key research by Ulrich (1984) could be used to influence the design of hospitals. [10]

- (b)* Discuss the validity of research into psychological effects of the built environment. [15]

- (c)* Jon works in the planning department of his local town council. He has been given the job of designing a major expansion to the town. His aim is to design the new part of town so that it has a positive effect on the health/wellbeing of the people who will live there.

What might a psychologist suggest about how Jon can achieve his aim? [10]

OPTION 4**Sport and exercise psychology**

- 8 (a)* Use the key research by Munroe-Chandler et al. (2008) to explain the benefits of imagery in sport. [10]

- (b)* Discuss the reliability of research into motivation in sport. [15]

- (c)* Debbie is the coach for her daughter's rugby team. It is halfway through the season and they have lost every match they've played. Some of the girls in the team are beginning to lose interest in the sport.

What advice might a sports psychologist give Debbie about how to motivate the players in her daughter's rugby team? [10]

END OF QUESTION PAPER

OCR

Oxford Cambridge and RSA

Copyright Information

OCR is committed to seeking permission to reproduce all third-party content that it uses in its assessment materials. OCR has attempted to identify and contact all copyright holders whose work is used in this paper. To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced in the OCR Copyright Acknowledgements Booklet. This is produced for each series of examinations and is freely available to download from our public website (www.ocr.org.uk) after the live examination series.

If OCR has unwittingly failed to correctly acknowledge or clear any third-party content in this assessment material, OCR will be happy to correct its mistake at the earliest possible opportunity.

For queries or further information please contact the Copyright Team, First Floor, 9 Hills Road, Cambridge CB2 1GE.

OCR is part of the Cambridge Assessment Group; Cambridge Assessment is the brand name of University of Cambridge Local Examinations Syndicate (UCLES), which is itself a department of the University of Cambridge.

Write the question number in the margin. Continue on the next page if needed.

Candidate name: *StemUp Model Solution*

- 1(a) *A non-biological treatment of a specific phobia, such as arachnophobia, is systematic desensitisation, a behaviourist therapy based on classical conditioning. The therapist first teaches the client deep muscle relaxation, breathing slowly and progressively tensing and releasing muscle groups, because relaxation is incompatible with fear (reciprocal inhibition). Client and therapist then build an anxiety hierarchy for the spider phobia, starting with the least frightening item, such as saying the word spider, moving through looking at a cartoon spider, then a photograph, a spider in a jar across the room, and ending with the most feared item, allowing a spider to walk on the hand. In weekly sessions the client is exposed to each step, in imagination first and then in vivo, while maintaining relaxation, and only moves up the hierarchy once the current step causes no anxiety. Over around six to twelve sessions the fear response is extinguished and replaced by a learned relaxation response (counterconditioning), so the spider no longer produces the phobic reaction.*
- 1(b) *One strength of systematic desensitisation for a specific phobia is its efficacy and speed. Because a phobia of spiders is a single, clearly defined conditioned fear, the graded hierarchy can target it directly, and outcome research such as McGrath et al. (1990) found that around 75 per cent of phobic clients respond to the technique, with gains usually achieved within a few months. The client is also in control of the pace of exposure and remains relaxed at each step, so unlike flooding it is rarely experienced as traumatic and drop-out is low, which makes it a more ethical and acceptable behavioural treatment.*
- One weakness is that it treats the symptoms rather than any underlying cause. The behaviourist model assumes the phobia was learned through association, so it only unlearns the conditioned response to spiders and does not address cognitions or other anxiety the person may have, meaning fear can return or be displaced on to a new object (symptom substitution). It also depends on the client being able to relax and to imagine the feared stimulus vividly, and it is impractical for complex disorders such as social phobia or agoraphobia, where the feared situations are varied and hard to recreate in a graded, controlled way, so its success is largely limited to simple specific phobias.*

Write the question number in the margin. Continue on the next page if needed.

2(a) Szasz argues that an illness is a physical lesion or disease of the body, so if a mental illness really had a bodily cause it would be a disease of the brain and belong to neurology, not psychiatry. Conditions once called mental illnesses, such as neurosyphilis and epilepsy, turned out to be bodily diseases, so the label mental illness is only used where no physical pathology can be found. He also argues that the mind is not an organ and so cannot literally be ill; what we call mental illness is really problems in living, behaviour that other people judge to be bad, deviant or disturbing. Finally, he says mental illness is a social, legal and political construct rather than a medical fact. It is used by the state, through involuntary commitment and the insanity defence, to control people whose behaviour society disapproves of, which shows it reflects society's judgements rather than any disease.

2(b) One line of argument supporting Szasz is that what counts as mental illness changes with time and culture, which suggests it is a social judgement rather than a disease. Homosexuality was a mental disorder in DSM-II until 1973 and was removed by a vote of the American Psychiatric Association, not by a medical discovery; drapetomania, the supposed illness of enslaved people who ran away, and culture-bound syndromes such as koro also show that diagnosis follows social values. Rosenhan's (1973) pseudopatients being admitted and labelled schizophrenic on the basis of one reported symptom supports the idea that the label is unreliable and socially constructed.

However, there is strong evidence against Szasz. Some disorders do have demonstrable biological bases, such as the high heritability of schizophrenia shown by Gottesman's concordance rates (48 per cent for monozygotic twins), the dopamine hypothesis and the effectiveness of antipsychotic drugs, which suggests they are genuine brain diseases even if the lesion is not yet fully understood. Furthermore, some disturbances, such as severe depression with suicidal intent or the hallucinations of psychosis, would cause concern in any culture at any time, so they are not simply society disapproving of odd behaviour. Finally, denying that mental illness exists has practical costs: diagnosis gives people access to treatment, support and legal protection, whereas

Write the question number in the margin. Continue on the next page if needed.

Szasz's view risks blaming individuals for problems in living that they cannot control. Overall Szasz usefully reminds us that diagnosis involves value judgements, but the claim that there is no such thing as mental illness goes too far.

- 3 Rosenhan's study contributes to understanding individual diversity because it showed that people differ in how they judge the same behaviour. In study one, no member of staff detected the eight pseudopatients, yet 35 of the 118 real patients voiced suspicions, saying things like "you're not crazy, you're a journalist", so patients and professionals read the same evidence differently. In study two, where a hospital was warned that pseudopatients would try to gain admission over three months, staff varied individually in their judgements of the 193 genuine patients: 41 were confidently judged to be pseudopatients by at least one staff member, 23 by a psychiatrist and 19 by both, showing diagnosis depends on the individual judge.

It contributes to understanding social diversity because the pseudopatients were treated differently simply because of their social role as mental patients. When they asked staff a polite question such as "when am I likely to be discharged?", 71 per cent of psychiatrists and 88 per cent of nurses moved on with head averted, whereas when a young woman asked a similar innocuous question of staff at Stanford University, every one stopped and answered. The label depersonalised and powerlessness followed, with over 2100 pills handed out and staff spending very little time with patients.

Finally it illuminates cultural diversity because it describes how, in the American culture of the early 1970s, psychiatry over-diagnosed, labelling all twelve admissions as schizophrenia (one as manic depression) and discharging them "in remission", and how patients in that culture's hospitals were treated with little respect, which may not be true of all cultures or periods.

- 4 A deterministic explanation is one which says that behaviour is caused by factors outside the person's control, so that given those causes the illness will follow and free will plays no part. Hard determinism sees mental illness as fully caused by prior factors, whereas soft determinism

Write the question number in the margin. Continue on the next page if needed.

allows that causes constrain but do not completely decide a person's behaviour.

Biological explanations of mental illness are strongly determinist. The genetic explanation of schizophrenia argues that the disorder is inherited, supported by Gottesman's concordance rates of 48 per cent for monozygotic twins compared with 17 per cent for dizygotic twins, and the dopamine hypothesis argues that excess dopamine activity causes positive symptoms, a view supported by the effect of antipsychotics that block D2 receptors. If a person's genes or neurochemistry decide whether they become ill, then the individual has no choice in the matter. However, the same evidence undermines complete determinism, because concordance is far from 100 per cent; roughly half of identical twins of a sufferer do not develop schizophrenia despite sharing all their genes, so biology creates a predisposition rather than a certainty. Much of the evidence is also correlational, for example raised dopamine levels may be a consequence of the illness or of medication rather than its cause, and it is very difficult to isolate a single cause in a disorder that is multi-factorial.

Behaviourist explanations are also determinist, treating phobias as conditioned responses to the environment, as in Watson and Rayner's Little Albert, who learned to fear a white rat through pairing with a loud noise. Here the person is a passive product of their learning history. Yet many people have frightening experiences without developing a phobia, and individual differences in response to the same environment suggest the person's own appraisal matters, which is hard to explain in purely deterministic terms.

Cognitive explanations such as Beck's cognitive triad and Ellis's irrational beliefs lie closer to soft determinism. Depressive thinking is seen as caused by early experience and faulty schemas, but because thoughts are seen as changeable the person is given some agency, and the success of cognitive behavioural therapy, which requires the client actively to challenge and choose different thoughts, suggests the illness is not completely determined. On the other hand, if CBT works by changing cognitions, the recovery is still being caused by a new set of thoughts, so even here there is an element of determinism.

Write the question number in the margin. Continue on the next page if needed.

The psychodynamic explanation is perhaps the most determinist of all, since Freud claimed adult neuroses were fixed by unconscious conflicts and childhood experiences, as in Little Hans, over which the adult has no control. Finally the diathesis-stress model combines a biological vulnerability with environmental triggers, and because it shows the same genotype producing different outcomes depending on experience, it suggests determinism is probabilistic rather than absolute.

In conclusion, explanations of mental illness are determinist to a large extent, especially the biological and behaviourist ones, which is scientifically useful because identifying causes allows treatment and removes blame from the sufferer. However, the imperfect concordance rates, the correlational nature of much evidence, individual differences and the effectiveness of talking therapies that rely on patient choice show that no single explanation can claim mental illness is fully determined, and a soft determinist, interactionist position is the most defensible.

5(a) Van Leeuwen et al. (2008) carried out a twin-family study of general intelligence in the Netherlands. The sample was 112 twin families, each with a pair of 9-year-old twins (both monozygotic and dizygotic), an older sibling where possible and both parents, giving a large family design that allowed the researchers to separate genetic from environmental influences. Children completed Raven's Standard Progressive Matrices and parents completed the Advanced Progressive Matrices, which are non-verbal tests of reasoning that minimise cultural and educational bias. Zygosity was confirmed by DNA testing, and structural equation modelling compared the correlations between MZ twins, DZ twins, siblings and parents to estimate the contribution of genes, shared environment and non-shared environment.

The key finding was that genetic factors accounted for 67 per cent of the variation in children's Raven scores. The correlation in IQ between MZ twins was much higher than between DZ twins and ordinary siblings, which is only explicable through heredity, because monozygotic twins share all or nearly all of their DNA whereas dizygotic twins and siblings share on average half,

Write the question number in the margin. Continue on the next page if needed.

yet all grew up in the same home. The study found no evidence of cultural transmission, meaning parents did not pass on intelligence by the way they brought up their children, and the correlation between the parents' own scores (phenotypic assortment, about 0.33) showed that intelligent parents tend to have children together and so pass on more of the relevant genes. The researchers also found a gene-environment interaction: for children with a genetic predisposition towards lower intelligence the environment mattered more, and this interaction explained an additional 9 per cent of variance, whereas children with high genetic potential were less affected by environment.

This research therefore explains that intelligence is largely a biological characteristic. Inherited genes produce the brain structures and processing speed that allow abstract reasoning, so children's scores resemble their biological relatives in proportion to the genes shared, and a child's genotype also sets how far a good or poor environment can push their intelligence up or down. The remaining third of the variance, and the interaction effect, show that biology is the dominant but not the only factor.

5(b) Research into intelligence raises many methodological issues, meaning problems with how intelligence is defined, measured and investigated that affect the validity and reliability of findings. Validity concerns whether a test or study really measures intelligence, while reliability concerns the consistency of the measurement.

The most fundamental issue is construct validity: there is no agreed definition of intelligence. Spearman argued for a single general factor g , which is what Raven's Progressive Matrices and the Wechsler scales are designed to tap, whereas Gardner's theory of multiple intelligences proposes at least eight independent intelligences including musical, bodily-kinaesthetic and interpersonal, and Goleman argued emotional intelligence should be included. If Gardner is right then Van Leeuwen et al. (2008) only measured one narrow form of intelligence with the Raven test and their claim that 67 per cent of variation in intelligence is genetic applies only to abstract reasoning. On the other hand, the Raven test is deliberately non-verbal and correlates

Write the question number in the margin. Continue on the next page if needed.

strongly with other IQ measures, which supports its validity as a measure of g , so the issue depends on which construct one accepts.

A related issue is cultural bias. Traditional IQ tests such as the Stanford-Binet were standardised on white Western populations and contain items that require culturally specific knowledge, so children from other cultures or lower socio-economic backgrounds score lower for reasons unconnected to ability. This is why Van Leeuwen chose the Raven matrices, which use abstract patterns, yet even these assume familiarity with pen and paper tests and a motivation to perform, so no test is entirely culture-fair and comparisons across groups may lack validity.

Reliability is also an issue. A good intelligence test must have high internal reliability, checked by split-half methods, and test-retest reliability, which the Wechsler scales achieve with correlations above 0.9. However, performance on any one day is affected by fatigue, anxiety and practice, and the Flynn effect, the rise of average IQ scores by about three points a decade, shows that scores are not fixed properties of individuals but change with environment, which questions what a single score tells us.

Research designs raise further issues. Twin and adoption studies such as Van Leeuwen's rely on the equal environments assumption, that MZ and DZ twins experience equally similar environments, but identical twins are often treated more alike, which inflates heritability estimates. Samples are often unrepresentative; Van Leeuwen's 112 Dutch families were volunteers from a twin register and the children were all aged nine, so generalising to other ages and cultures is risky. Heritability estimates are also population-specific statistics, not statements about individuals, and the gene-environment interaction found in the study shows that genes and environment cannot cleanly be separated, so cause and effect is difficult to establish from correlational family data. Brain-imaging research linking intelligence to brain size or processing speed offers more objective biological measures, but it is also correlational and uses small samples.

Write the question number in the margin. Continue on the next page if needed.

In defence of this research, large samples, DNA zygoty testing, standardised tests and replication across many countries have produced consistent heritability estimates of 50 to 80 per cent, which gives the findings considerable reliability. In conclusion, intelligence research is methodologically sound in its reliability but remains vulnerable on validity, because the central construct is contested and culturally loaded, so findings should be interpreted as describing performance on particular tests in particular populations rather than intelligence in any absolute sense.

5(c) A psychologist would first advise Nasra's company to decide what kind of intelligence matters for designing computer games and to measure that directly, because a test is only useful if it is valid for the job. Game design requires abstract reasoning, spatial visualisation, problem solving and creativity rather than general knowledge, so the psychologist would recommend a non-verbal test of fluid reasoning such as Raven's Advanced Progressive Matrices as the core measure. This presents applicants with patterns of increasing difficulty and asks them to pick the missing piece, which closely resembles the logical problem solving used when designing game mechanics and debugging levels. It is also culture-fair, so applicants from different countries and language backgrounds, common in the games industry, are not disadvantaged, and it has well-established reliability and norms so Nasra can compare each candidate's score against the adult population.

Because a single score for g may be too narrow, the psychologist would suggest supplementing it with subtests from the Wechsler Adult Intelligence Scale that are relevant to the role, such as block design and matrix reasoning for spatial ability and digit span for working memory, which a designer needs to hold many interacting rules in mind. Drawing on Gardner's multiple intelligences, the company might also include a short practical task, for example asking applicants to design a simple game level or solve a puzzle in a prototype, which assesses spatial and logical-mathematical intelligence in a realistic way and has high predictive validity. If the designers work in teams, a measure of emotional intelligence or interpersonal skills could be added, since collaboration is part of the job.

Write the question number in the margin. Continue on the next page if needed.

The psychologist would advise administering the tests under standardised conditions, ideally online with a fixed time limit so that all applicants are treated the same and the results are reliable, and would warn Nasra that tests should be used to shortlist alongside portfolios rather than as the sole decision, since IQ tests measure potential rather than acquired skill and can be affected by test anxiety and practice. Used this way, the combination of a Raven-style reasoning test, job-relevant WAIS subtests and a practical design task is the best method because it is valid for the work, fair to a diverse applicant pool, quick to administer to large numbers and produces quantitative data that can be compared objectively.

6(a) Dixon et al. (2002) investigated whether a defendant's regional accent, race and the type of crime affected attributions of guilt. The participants were 119 white undergraduate psychology students at University College Worcester, none of whom spoke with a Birmingham accent. In an independent measures design they listened to a two-minute recording of a supposed police interview in which a male suspect denied involvement in a crime. The same actor, a code-switcher, recorded the suspect's part in either a Birmingham (Brummie) accent or a standard, received pronunciation accent. The transcript also varied the suspect's race, described as black or white, and the type of crime, either blue-collar (armed robbery) or white-collar (cheque fraud). Participants then rated the suspect's guilt on a seven-point scale from innocent to guilty, and completed the Speech Evaluation Instrument, rating the speaker on superiority, attractiveness and dynamism.

The main finding was that the Brummie-accented suspect was rated significantly more guilty than the standard-accented suspect. There was also a significant three-way interaction: guilt ratings were highest of all when the suspect was black, had a Brummie accent and was accused of the blue-collar crime of armed robbery. The Brummie speaker was also rated as lower in superiority (less intelligent, less competent), and this rating partly mediated the effect of accent on guilt, although attractiveness and dynamism did not.

This shows how the characteristics of a defendant can affect whether they are found guilty.

Write the question number in the margin. Continue on the next page if needed.

Jurors hold stereotypes about social groups, so a non-standard regional accent acts as a cue to lower status and competence, which is then associated with criminality, and a black defendant accused of a stereotypically "black" violent crime is judged more harshly because the crime fits the juror's schema. Even though the evidence is identical, these extra-evidential factors bias the interpretation of the evidence and make a guilty verdict more likely for defendants who sound or look like the stereotype of a criminal. However, the study used mock jurors judging guilt alone from a recording rather than a real jury deliberating on a full trial, so Dixon et al. show how such characteristics could influence verdicts rather than proving they decide real cases.

6(b) Ethical considerations are the standards set out in the BPS code, such as informed consent, the right to withdraw, deception, protection from harm, confidentiality and debriefing, which researchers must weigh against the value of their findings. Research into psychology and the courtroom, which examines how juries are persuaded by features of defendants, witnesses and evidence, raises these issues in distinctive ways.

Much of the research uses mock juries, which can be ethically sound. In Dixon et al. (2002) the 119 students gave consent to listen to a recorded interview and rate guilt, no one was harmed and participants could withdraw at any time, so the study stayed within the guidelines. However, mock jury research necessarily involves a degree of deception, since the suspect was an actor and participants may have believed the case was real, and the race and accent manipulation could be criticised as reinforcing stereotypes or causing participants discomfort when debriefed about their own prejudice. Similarly, Castellow et al. (1990) asked students to judge an attractive or unattractive defendant in a sexual harassment case, and Penrod and Cutler (1995) manipulated witness confidence; these designs expose participants to potentially distressing material about crime, which raises issues of protection from harm, although the harm is small and debriefing reduces it. Such studies can also be defended because their artificiality protects real defendants: it would be clearly unethical to manipulate a real trial.

The more serious issue is that real jury research is largely impossible, because in the UK the

Write the question number in the margin. Continue on the next page if needed.

Contempt of Court Act makes it illegal to ask jurors about deliberations, protecting confidentiality and the defendant's right to a fair trial. This forces psychologists into less valid shadow jury or simulation designs, which is a case where ethics reduces validity. Where researchers have observed real juries, as in the Chicago Jury Project when deliberations were secretly recorded, the invasion of privacy and lack of consent caused a scandal and led to laws banning the practice.

A broader ethical consideration is the use to which findings are put. Research showing that a confident witness is believed, that an attractive or standard-accented defendant seems less guilty, or that the story model (Pennington and Hastie 1988) means jurors are persuaded by evidence presented in story order, can be used by lawyers to coach witnesses and present cases in whatever way wins. This could help a guilty person be found innocent or an innocent person be found guilty, which is socially sensitive, and raises the question of whether the end of understanding justice justifies the means of giving advocates tools to manipulate juries. On the other hand, the same research allows the courts to protect defendants, for example by giving judges' instructions about witness confidence or screening defendants' appearance, so the knowledge arguably does more good than harm.

In conclusion, courtroom research is generally low-risk to its participants and mostly satisfies consent and harm guidelines, but it faces unavoidable tensions: confidentiality and fair trial rights prevent studying real juries, deception is built into mock trials, and the findings are socially sensitive because they can be used to pervert as well as improve justice. Psychologists should therefore prioritise thorough debriefing, avoid stereotyping materials where possible, and publish findings in ways that inform court procedure rather than only advocates.

- 6(c) A psychologist would tell Alan that wrong verdicts often result from jurors being influenced by factors that are not evidence, so the court system should remove or neutralise those factors. Dixon et al. (2002) showed that a defendant with a Brummie accent, especially a black defendant accused of armed robbery, was rated more guilty than a standard-accented defendant

Write the question number in the margin. Continue on the next page if needed.

on identical evidence, and Castellow et al. (1990) found attractive defendants were less likely to be found guilty. Alan could therefore allow defendants and vulnerable witnesses to give evidence from behind screens or by written statement read in a neutral voice, so that jurors cannot be swayed by appearance or accent, and he could require that any persuasion techniques such as expert witnesses or demonstrative evidence be available equally to prosecution and defence so neither side gains an unfair extra-evidential advantage.

Second, Alan should introduce juror training and judicial instructions. Penrod and Cutler (1995) found that jurors convicted in 67 per cent of cases when an eyewitness said she was 100 per cent confident compared with 60 per cent when she was 80 per cent confident, even though confidence is a poor predictor of accuracy. A short video before trials, or a standard direction from the judge, should warn jurors not to give undue weight to witness confidence and to be wary of the CSI effect, the expectation that forensic evidence will always be available and conclusive, which leads to acquittals when it is absent and over-conviction when it is presented.

Third, Alan could improve how evidence is presented. Pennington and Hastie's story model shows jurors construct a narrative, and evidence presented in story order was far more persuasive than witness order, so he could require both sides to present in chronological story order and provide jurors with written summaries, timelines and the right to take notes and ask questions, reducing the burden on memory in long trials.

Finally, the psychologist would address the deliberation process itself. Asch's research shows people conform to a majority, and Hastie et al. found that juries typically decide in the direction of the first ballot majority. Alan could therefore require a secret first ballot before discussion, so that each juror records an independent view and minority voices are not silenced, keep juries at twelve rather than smaller sizes to make a lone dissenter less likely, and give the jury foreperson guidance on evidence-driven rather than verdict-driven deliberation. Together these measures target the main sources of bias and should reduce the number of wrong verdicts.

Write the question number in the margin. Continue on the next page if needed.

7(a) Ulrich (1984) investigated whether a view of nature from a hospital window affected recovery from surgery. He used the records of 46 patients who had undergone cholecystectomy (gall bladder removal) at a suburban Pennsylvania hospital between 1972 and 1981. All had been on the same floor, where rooms on one side of the corridor looked out on to a small stand of deciduous trees and rooms on the other side faced a brown brick wall. Patients were matched in 23 pairs on age, sex, smoking, weight and the year of surgery, and only patients admitted between May and October, when the trees were in leaf, were included. A nurse blind to the hypothesis extracted data on length of post-operative stay, the number and strength of analgesics taken, nurses' notes and minor complications.

Ulrich found that patients with the tree view stayed in hospital for a shorter time after surgery (7.96 days compared with 8.70 days for the wall group), received far fewer negative comments in nurses' notes such as "upset and crying" or "needs much encouragement" (1.13 compared with 3.96 per patient), took fewer doses of moderate and strong painkillers between days two and five and more weak doses such as aspirin, and had slightly fewer minor post-surgical complications. He concluded that a natural view reduces stress and improves recovery, in line with the idea that nature has restorative effects.

This research could influence hospital design in several ways. Architects should orient as many patient rooms as possible towards gardens, trees or planted courtyards, and where a hospital is built on a tight urban site they should create internal green courtyards or roof gardens that wards overlook, rather than facing blank walls or car parks. Within a ward, beds should be positioned close to windows so that patients lying down can actually see the trees, with low sills and large glazing, and patients recovering from the most serious operations, who stand to benefit most, should be given priority for rooms with natural views. Where a natural view is impossible, Ulrich's later work suggests large nature murals or photographs can offer some benefit, so corridors and treatment rooms could display natural scenes. Finally, because shorter stays and reduced analgesic use save money, hospital managers could justify the extra cost of landscaping and planting trees around the building on economic as well as humane grounds.

Write the question number in the margin. Continue on the next page if needed.

7(b) Validity refers to whether a piece of research measures what it claims to measure and whether its conclusions are true. It includes internal validity (did the manipulated variable cause the effect), ecological validity (does the setting reflect real life), population validity (can the sample be generalised) and predictive validity (do findings forecast real outcomes). Research into the psychological effects of the built environment, such as housing design, hospital design, noise and density, is especially open to validity problems because people cannot easily be randomly allocated to buildings.

A major issue is isolating the variable. Ulrich (1984) compared patients with a tree view against a wall view and found shorter stays and less analgesic use, which has high ecological validity because the patients were real, having real surgery in a real hospital. However, it was a natural experiment using archival records, so other differences between the two sides of the corridor, such as daylight, noise from the car park or which nurses were on duty, cannot be ruled out, which threatens internal validity. Ulrich strengthened validity by matching 23 pairs on age, sex, smoking and year of surgery and by having the data extracted by a nurse blind to the hypothesis, but the sample of 46 gall bladder patients in one suburban Pennsylvania hospital limits population validity, since the effect might not apply to other illnesses, cultures or urban hospitals.

Research into high-rise housing and urban renewal faces similar problems. Newman's work on defensible space, illustrated by the failure of the Pruitt-Igoe estate in St Louis and the comparison of Van Dyke and Brownsville projects, linked building height and anonymous shared space to crime and vandalism, yet residents in such estates also differ in poverty, unemployment and social problems, so it is difficult to know whether the architecture or the social conditions caused the effects. Wilson and Kelling's broken windows theory has strong face validity and influenced policing, but tests of it such as Keizer et al.'s field experiments in Groningen show that disorder cues increase littering and theft, while the theory's claims about serious crime remain contested.

Write the question number in the margin. Continue on the next page if needed.

Much built-environment research relies on self-report, such as questionnaires on wellbeing or satisfaction with a neighbourhood after regeneration. These are vulnerable to demand characteristics and social desirability bias, because residents who have been given new homes may feel obliged to report feeling better, and to the problem that improvements in wellbeing may reflect the attention of the renewal process rather than the buildings. Objective measures, such as Ulrich's hospital records or crime statistics, avoid this but may be recorded inconsistently.

Laboratory research offers better control but weaker ecological validity. Glass and Singer's studies of noise and performance manipulated predictability of noise precisely, but short exposure to recorded noise in a laboratory may not reflect living beside a motorway for years. In contrast, field studies such as Bronzaft and McCarthy's comparison of reading scores in classrooms on the noisy and quiet sides of a school next to an elevated railway have high ecological validity, and a later study after acoustic treatment showed the difference disappeared, which gives the finding good predictive validity. Predictive validity is the ultimate test for this field: Ulrich's findings have been replicated and adopted in evidence-based hospital design, and improvements in patient outcomes after redesign suggest the original research was valid.

In conclusion, built-environment research gains ecological validity by studying real buildings and real residents but pays for it with reduced internal validity, confounding variables and often small or unrepresentative samples. Triangulating field studies with controlled experiments and objective measures, as Ulrich did, offers the most valid picture, and the fact that design changes based on this research have produced measurable improvements suggests that, despite its limitations, the research is reasonably valid.

7(c) A psychologist would advise Jon that the layout and features of the new district can directly affect residents' stress, health and sense of community, and would make several evidence-based suggestions. First, Jon should build in green space. Ulrich (1984) found that hospital patients

Write the question number in the margin. Continue on the next page if needed.

who could see trees recovered faster and needed fewer painkillers than those facing a wall, and Kaplan's attention restoration theory and Kuo and Sullivan's work on Chicago housing found that residents with views of trees and grass reported less aggression and mental fatigue. Jon should therefore plan parks, tree-lined streets and small pocket gardens so that every home is within a short walk of greenery and as many windows as possible overlook trees rather than roads or blank walls.

Second, Jon should avoid high-rise, high-density housing. Newman's concept of defensible space, and the failure of Pruitt-Igoe, show that anonymous shared spaces such as long corridors and large unsupervised grounds encourage crime and fear, while low-rise homes with clearly marked private gardens and semi-private shared courtyards give residents a sense of ownership and natural surveillance. Low-density housing arranged in short streets or cul-de-sacs with windows facing the street, and clear boundaries such as low fences and hedges, would increase territoriality and social contact and reduce crime and anxiety.

Third, the psychologist would stress noise reduction, since Glass and Singer and Bronzaft showed that uncontrollable noise impairs concentration, raises stress and lowers children's reading scores. Jon should route through-traffic around rather than through the residential area, use earth banks, planting and sound-insulating building standards near main roads, and keep schools and homes away from noisy industry.

Fourth, Jon should design for active travel and community. Providing segregated cycle lanes, safe footpaths and local facilities such as shops, a school, a health centre and community hall within walking distance reduces unnecessary car travel, pollution and sedentary behaviour, encourages exercise that protects mental and physical health, and increases chance encounters that build social networks. Festinger's work on friendship formation in housing shows that people become friends with those whose paths they cross regularly, so shared facilities, benches and playgrounds placed at natural junctions will strengthen community cohesion.

Write the question number in the margin. Continue on the next page if needed.

Finally, Jon should maintain the environment well, because broken windows theory shows that visible neglect such as litter and graffiti signals that no one cares and leads to further disorder and fear of crime. By combining green views, defensible low-density housing, quiet streets, walkable local facilities and good upkeep, Jon can design a district that has a measurable positive effect on residents' health and wellbeing.

8(a) Munroe-Chandler et al. (2008) investigated the relationship between imagery use and self-confidence and self-efficacy in young footballers. The sample was 122 male and female soccer players aged 11 to 14 from Ontario, Canada, divided into recreational (house league) players and competitive (travel team) players. Each completed three self-report questionnaires in a group setting: the Sport Imagery Questionnaire for Children (SIQ-C), which measures how often players use the five functions of imagery, namely cognitive specific (imaging skills), cognitive general (strategies), motivational specific (goals and outcomes), motivational general-arousal (controlling anxiety and excitement) and motivational general-mastery (MG-M, imaging being confident, focused and in control); the self-confidence subscale of the Competitive State Anxiety Inventory for children (CSAI-2C); and a soccer-specific self-efficacy questionnaire asking how confident they were that they could perform particular skills.

Regression analyses showed that of the five imagery types only MG-M imagery was a significant predictor of both self-confidence and self-efficacy, and this was true for both recreational and competitive players. MG-M imagery accounted for more of the variance in confidence among recreational players than competitive players, although the difference was not statistically significant. The researchers concluded that imagining oneself coping and mastering challenging situations in sport is linked to higher confidence in young athletes regardless of their level.

This research explains several benefits of imagery in sport. First, imagery that focuses on mastery, such as a player picturing herself staying composed after a missed shot or confidently beating a defender, builds self-confidence, which Vealey's model shows leads to better

Write the question number in the margin. Continue on the next page if needed.

performance and persistence. Second, it raises self-efficacy, the belief that one can execute specific skills, which Bandura argues determines how much effort athletes invest and how long they persist when they fail; imagery acts like vicarious or enacted experience, giving the brain a successful rehearsal without the cost of physical failure. Third, because the effect was found in recreational as well as competitive players, imagery is a cheap, accessible mental skill that coaches can teach to youngsters at all levels, perhaps most usefully to beginners whose confidence is fragile. Coaches should therefore build MG-M imagery scripts into training, asking players to spend a few minutes before practice imagining themselves performing confidently under pressure. One caution is that the study was correlational, so confident players may simply use more imagery, but combined with experimental evidence the findings strongly support imagery as a tool for improving confidence and performance.

8(b) Reliability refers to the consistency of a measure or a set of findings: a reliable questionnaire gives the same results when repeated (test-retest) and its items agree with one another (internal consistency), and reliable research produces the same findings when replicated by other researchers with other samples. Research into motivation in sport, including self-efficacy, sport confidence, competitiveness and imagery, relies heavily on psychometric questionnaires, so reliability is central to judging it.

Much of this research scores well on reliability because it uses standardised, validated instruments. Munroe-Chandler et al. (2008) measured imagery use with the Sport Imagery Questionnaire for Children, self-confidence with the CSAI-2C and self-efficacy with a specially constructed scale, all administered under the same group conditions to 122 players, and the SIQ-C had been developed from the adult SIQ with established internal consistency (Cronbach's alpha values typically above 0.7 for each subscale). Using several items to measure each construct, so that a single misread question does not distort the score, increases reliability. Similarly, Gill and Deeter's Sport Orientation Questionnaire, which measures competitiveness, win orientation and goal orientation, was developed through repeated administrations and showed high test-retest reliability over several weeks, and Vealey's Trait and State Sport-Confidence

Write the question number in the margin. Continue on the next page if needed.

Inventories were likewise refined to produce consistent scores. Large samples such as Munroe-Chandler's reduce the influence of individual quirks and make findings more stable.

Reliability is further supported by replication. The finding that MG-M imagery predicts confidence has been found in adult athletes by Moritz et al. and Callow and Hardy, and Bandura's self-efficacy theory has been supported across many sports and age groups, so the pattern is consistent across studies. Standardised procedures mean other researchers can repeat the research exactly.

However, there are threats to reliability. Motivation is a state as well as a trait, so a player's responses on a self-confidence questionnaire may vary from day to day depending on recent results, selection or mood, meaning test-retest reliability for state measures is inherently limited and a single administration, as in Munroe-Chandler et al., gives only a snapshot. Self-report questionnaires with children aged 11 to 14 raise further concerns: young players may interpret items about imagery differently, respond in a socially desirable way to please coaches, or answer carelessly in a group setting, which produces inconsistency. Correlational designs are also only as reliable as the measures within them, and when self-efficacy is measured with sport-specific scales constructed for a single study, as with the soccer self-efficacy questionnaire, its reliability is not established before use. In addition, cross-cultural replication is limited; much of the research comes from North America, and motivation constructs such as competitiveness may be understood differently elsewhere, so findings may not replicate consistently across cultures.

In conclusion, research into motivation in sport is reasonably reliable because it relies on standardised multi-item questionnaires with demonstrated internal consistency, large samples and repeated replication of key findings such as the link between MG-M imagery and confidence. Its main weakness is that motivation fluctuates and is measured by self-report, so results are consistent in their broad patterns but less consistent at the level of the individual athlete, and coaches should treat questionnaire scores as indicative rather than fixed.

Write the question number in the margin. Continue on the next page if needed.

8(c) A sports psychologist would tell Debbie that after losing every match the girls' self-confidence and self-efficacy are low, and that motivation can be rebuilt by changing what success means to the team and how they picture themselves. First, the psychologist would recommend motivational general-mastery imagery. Munroe-Chandler et al. (2008) found that MG-M imagery was the strongest predictor of self-confidence and self-efficacy in 11 to 14 year old footballers, particularly recreational ones. Debbie could spend five minutes at the start of each training session leading a guided imagery script in which each girl closes her eyes and imagines herself making a strong tackle, staying calm after conceding a try, and feeling confident and in control; players could also do this the night before matches.

Second, Debbie should rebuild self-efficacy using Bandura's four sources. Performance accomplishments are the most powerful, so training should be structured so that players succeed at achievable, progressively harder drills and are reminded of specific things they did well in matches, such as winning more scrums than last week. Vicarious experience can come from watching video of a similar girls' team or older players executing skills, verbal persuasion from Debbie giving specific, genuine praise rather than empty encouragement, and the girls can be taught to interpret pre-match nerves as readiness rather than fear.

Third, the psychologist would suggest shifting from outcome goals to process and performance goals. Because the team cannot control whether they win, Debbie should set SMART goals the players can control, such as completing 80 per cent of tackles or scoring one try per half, and review them after each match so that the girls experience success even in a defeat. Drawing on Gill and Deeter's sport orientation research, Debbie should encourage a goal orientation rather than a win orientation, and from self-determination theory she should build intrinsic motivation by giving the girls choice over some drills and positions (autonomy), feedback that shows skill improvement (competence) and team-building activities that create belonging (relatedness). Praise and small rewards should be for effort and improvement, since rewards for winning alone would undermine intrinsic motivation.

Write the question number in the margin. Continue on the next page if needed.

Finally, Debbie could help the players attribute defeats to controllable, unstable factors such as effort and tactics rather than to lack of ability, so that losing does not become learned helplessness, and could set a mastery climate in which mistakes are treated as learning. Together these strategies should restore the girls' confidence and enjoyment and keep them in the sport for the rest of the season.