



Oxford Cambridge and RSA

Monday 9 June 2025 – Morning

A Level Psychology

H567/03 Applied psychology

Time allowed: 2 hours



You must have:

- the OCR 12-page Answer Booklet

INSTRUCTIONS

- Use black ink.
- Write your answer to each question in the Answer Booklet. The question numbers must be clearly shown.
- Fill in the boxes on the front of the Answer Booklet.
- Answer **all** the questions in Section A.
- Choose **two** options in Section B and answer **all** the parts of the question for each option.

INFORMATION

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

ADVICE

- Try to answer every part of each question you choose.
- Read each question carefully before you start your answer.

Section A

Issues in mental health

1 Outline the biochemical explanation of mental illness. [5]

2
(a) Describe **one** way mental disorders can be categorised. [4]

(b) Outline **one** problem a clinical psychologist might have when categorising mental disorders. Use an example of a mental disorder to support your answer. [4]

3 As part of the key research by Rosenhan (1973), the pseudopatients approached members of staff within the hospitals with a polite request for relevant information. The table contains details of how hospital staff responded to these requests. The table also contains results from when a person approached members of teaching staff at a university with a similarly polite request for relevant information.

	Percentage of hospital staff in a psychiatric hospital doing each action (based on 1283 attempted interactions)	Percentage of teaching staff at a university doing each action (based on 14 attempted interactions)
Moves on, head averted	88	0
Makes eye contact	10	0
Pauses and chats	2	0
Stops and talks	0.5	100

(a) Outline **one** conclusion that can be drawn from these findings. [3]

(b) Outline **one** suggestion a psychologist might make about how these findings could be used. [3]

(c) Assess the reliability of these findings. [6]

4* Discuss the validity of definitions of abnormality. [10]

Section B**Options**

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

OPTION 1**Child psychology****5**

(a)* Outline the key research by Gibson and Walk (1960) **and** use it to show the extent to which depth can be perceived by infants. **[10]**

(b)* Discuss the reliability of research into perceptual development. **[15]**

(c)* Nina is the manager of a pre-school playgroup for children aged 2–5 years. She is concerned that young children have too much ‘screen time’ (i.e. playing on electronic devices) so wants to use play to aid their development.

Outline at least one play strategy a psychologist might suggest to Nina to use with the children at her playgroup to help develop their perception. **[10]**

OPTION 2**Criminal psychology****6**

(a)* Outline the key research by Hall and Player (2008) **and** use it to show the extent to which bias affects the collection and processing of forensic evidence. **[10]**

(b)* Discuss to what extent research into the collection and processing of forensic evidence is scientific. **[15]**

(c)* From a crime scene, the police recover different examples of forensic evidence including fingerprints and a handwritten note. Experts will analyse the forensic evidence to help the police identify a suspect.

Outline at least one strategy a psychologist might suggest to the experts to reduce bias in the collection and processing of forensic evidence. **[10]**

OPTION 3**Environmental psychology****7**

- (a)* Outline the key research by Ulrich (1984) **and** use it to show the extent to which the view through a window may influence recovery from surgery. **[10]**
- (b)* Discuss the usefulness of research into the psychological effects of the built environment. **[15]**
- (c)* Sam has the job of converting an office block into apartments for people to live in. The office block is in the town centre and is twelve storeys high. Sam wants to design apartments that will improve the health/wellbeing of the people choosing to live there.

Outline at least one strategy a psychologist might suggest for how Sam can design apartments that will improve the health/wellbeing of the people choosing to live there. **[10]**

OPTION 4**Sport and exercise psychology****8**

- (a)* Outline the key research by Lewis et al (2014) **and** use it to show the extent to which dance can improve mental health. **[10]**
- (b)* Discuss the validity of research into exercise and mental health. **[15]**
- (c)* Amit is a primary school teacher with a class of children in Year 6 (aged 10–11 years). The children will soon be taking tests, and Amit wants to get them doing something that will improve their mental health. The school is near open countryside.

Outline at least one strategy a psychologist might suggest to Amit for how exercise could be used to improve the mental health of the school pupils. **[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 OCR Copyright Team, The Triangle Building, Shaftesbury Road, Cambridge CB2 8EA.

OCR is part of Cambridge University Press & Assessment, 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 The biochemical explanation says that mental illness is caused by an imbalance of neurotransmitters, the chemicals that carry messages across the synaptic gap between neurons. The most well known example is the monoamine hypothesis of depression, which states that depression is caused by too little serotonin (and noradrenaline) in the brain. Serotonin normally regulates mood, sleep and appetite and keeps the brain in a stable state, so when too little is released into the synapse the next neuron is not excited enough and the person experiences low mood, poor sleep and loss of appetite. A second example is the dopamine hypothesis of schizophrenia, where too much dopamine, or too many sensitive D2 receptors, leads to over-stimulation of neurons in the mesolimbic pathway and produces positive symptoms such as hallucinations and delusions.
- 2(a) One way mental disorders can be categorised is using the Diagnostic and Statistical Manual of Mental Disorders, now in its fifth edition (DSM-5). It is published by the American Psychiatric Association and is used mainly by psychiatrists and clinical psychologists in the USA to give a diagnosis. The manual is divided into chapters or categories of disorder such as schizophrenia spectrum and other psychotic disorders, depressive disorders, anxiety disorders and obsessive-compulsive disorders. Within each category the DSM lists the specific symptoms of each disorder, how many of them must be present and for how long. For example, for major depressive disorder a patient must show at least five of nine symptoms, including depressed mood or loss of interest, nearly every day for at least two weeks.
- 2(b) One problem a clinical psychologist has is that they usually only have the patient's own self-report of their symptoms to go on. They have to compare what the patient tells them with the lists of symptoms in a tool such as the DSM or ICD, but the patient may forget symptoms, describe them badly, exaggerate them or even lie about them, so the diagnosis is only as valid as the report. For example, in Rosenhan's study eight healthy pseudopatients simply reported hearing a voice saying 'empty', 'hollow' and 'thud'. Every one of them was

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

Candidate name: *StemUp Model Solution*

admitted to a psychiatric hospital, and seven were diagnosed with schizophrenia on the basis of this single false report. This shows that a psychologist can easily misdiagnose someone when they cannot check whether the reported symptoms are real.

3(a) One conclusion is that people who are labelled as patients in a psychiatric hospital are largely ignored by staff, whereas the same polite request in an ordinary setting is answered. In the hospital 88% of staff moved on with their head averted and only 0.5% stopped and talked, but at the university 100% of the teaching staff stopped and talked to the person asking. A plausible reason is that the psychiatric label depersonalises the patient, so staff see them as unimportant and not worth their time, while a person on a university campus is treated as a normal, valued individual.

3(b) A psychologist might suggest that the hospital trains its nurses and attendants to stop and talk to patients when they make a polite request for information, instead of walking past with their head averted. The training could be delivered by a clinical psychologist and could use role play of common requests such as 'When am I likely to be discharged?', with staff practising giving eye contact and a proper answer. This is needed because the findings show that in the hospital only 0.5% of staff stopped and talked and 88% moved on, which shows that patients are being depersonalised, whereas at the university every member of staff stopped and talked, so the behaviour can clearly be changed.

3(c) Reliability refers to how consistent the findings are. These findings appear quite reliable for the hospital condition because they are based on 1283 attempted interactions across several hospitals, and a very clear, consistent pattern emerged, with 88% of staff moving on and head averted and only 0.5% stopping to talk. Such a large number of trials means a few unusual staff members could not distort the overall trend, so if the pseudopatients had repeated the requests we would expect very similar percentages. The findings for the university

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

Candidate name: *StemUp Model Solution*

condition are far less reliable, however, because they rest on just 14 interactions with one person at one university. With so few trials a couple of unusually helpful lecturers could produce the 100% figure, so the result might not be consistent on replication. A further issue is that the requests were made by different pseudopatients in different hospitals, so they may not have been fully standardised; differences in tone of voice, wording or appearance could have changed how staff responded, lowering the inter-observer consistency. Finally, it is very hard to replicate the study to check its reliability, because it needs researchers to gain admission to psychiatric hospitals by faking symptoms. Overall the hospital findings are reliable but the comparison with the university is based on too few trials to be fully trusted.

4 Validity refers to whether a definition of abnormality measures what it claims to, in other words whether it accurately identifies people who really have a mental disorder and excludes those who do not.

Statistical infrequency defines abnormal behaviour as behaviour that is rare, falling more than two standard deviations from the mean on a normal distribution, such as an IQ below 70 for intellectual disability. This has some validity because it uses objective, quantitative data and does not rely on the clinician's opinion. However its validity is limited because many statistically rare behaviours are not disorders, for example an IQ of 140 or exceptional musical ability, while some disorders are so common that they are not statistically infrequent at all; depression affects about one in six adults in the UK. Statistics are also affected by culture, because in some cultures symptoms that are frowned upon are not reported or diagnosed, so the figures underestimate how frequent a disorder really is.

Deviation from social norms defines abnormality as behaviour that breaks the unwritten rules of a society. This lacks validity because norms change over time and place, so the same behaviour can be classed as a disorder in one era and not another. Homosexuality was removed from the DSM in 1973 but is still seen as abnormal in some parts of the world, which suggests the definition measures social acceptability rather than mental illness. It can also be

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

Candidate name: *StemUp Model Solution*

used to control people who are simply different, such as political dissidents.

Deviation from ideal mental health (Jahoda) lists criteria such as self-actualisation, accurate perception of reality and autonomy. The problem is that these are vague and hard to operationalise, and almost nobody meets all of them, so the definition over-diagnoses. Failure to function adequately is more useful because it considers the person's own distress, but it is still subjective: someone hearing voices who copes well would not be abnormal under this definition but would be under deviation from social norms.

In conclusion, the definitions that rely on subjective judgement (social norms, ideal mental health, failure to function) have low validity, and even the objective statistical definition cannot by itself decide what counts as a disorder, so clinicians need to combine definitions to diagnose accurately.

5(a) Gibson and Walk aimed to find out whether depth perception is innate or learned, by testing whether infants and young animals would avoid a visual drop as soon as they could move around. The sample was 36 human infants aged 6 to 14 months, each tested with their mother, plus a range of young animals including chicks, turtles, rats, lambs, kids, pigs, kittens and puppies. The method was a laboratory experiment using the visual cliff: a centre board laid across a large sheet of thick glass about a foot above the floor. On the shallow side a checked pattern was fixed directly under the glass, while on the deep side the same pattern was placed on the floor below, creating the appearance of a cliff although the glass meant it was perfectly safe. Each infant was placed on the centre board and the mother called the child first from the cliff side and then from the shallow side, and the researchers recorded which side the child crawled on. Animals were placed on the board and their choice of side was observed.

The results show that depth can be perceived to a great extent by the time infants can crawl. Of the 27 infants who moved off the board, all of them crawled onto the shallow side at least once, but only 3 ever crawled onto the deep side. Many infants cried or crawled away

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

Candidate name: *StemUp Model Solution*

from their mother when she called from the cliff side, showing they could see the drop even though they could feel the solid glass with their hands. Some infants did back onto the deep side by accident while setting off towards their mother across the shallow side, which suggests that the ability to perceive depth develops before the child has full control of their movement. The animal findings support this: day-old chicks, kids and lambs never stepped onto the deep side, and kittens avoided it as soon as they could walk at four weeks, suggesting depth perception is innate in species that need it from birth. Overall the study shows that infants can perceive depth by at least six months, although it cannot prove it is present at birth because the babies had six months of experience before testing.

5(b) Reliability refers to the consistency of research: whether a standardised procedure would produce the same results if repeated, and whether different observers agree about what they have seen. Research into perceptual development has a number of strengths in this respect but also some weaknesses.

A strength is that key studies use highly standardised, objective procedures. Gibson and Walk's visual cliff is a fixed piece of apparatus with the same glass, the same checked pattern and the same drop for every infant, and each mother called from the cliff side and then the shallow side in the same order. The dependent variable, which side the infant crawled onto, is clearly observable and gives quantitative data (27 crawled on the shallow side, only 3 on the deep side), so there is little room for observer disagreement. This means the study is easy to replicate exactly, and indeed the visual cliff has been repeated many times with similar findings, such as Campos et al finding heart-rate changes in younger infants placed over the deep side, which supports the reliability of the original.

Similarly, Blakemore and Cooper's study of kittens raised in vertical or horizontal striped cylinders used a controlled laboratory environment with a standardised exposure of five hours a day, and later tested the kittens' neurons with objective electrophysiological recordings. Because the environment and the measurement were tightly controlled, the study could be

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

Candidate name: *StemUp Model Solution*

replicated and the finding that no neuron responded to the orientation the kitten had never seen is consistent and dependable.

However, some observations in perceptual research are more subjective, particularly when the participants are non-human or pre-verbal. Blakemore and Cooper described the kittens as 'clumsy' and as bumping into objects, and Gibson and Walk noted that infants 'cried' or 'peered down'. These behavioural descriptions rely on the researcher's interpretation and no inter-rater reliability is reported, so another observer might have coded the behaviour differently. Fantz's preferential looking studies also depend on the observer judging the direction of a baby's gaze through a peephole, which is open to error and bias.

Sample size is another issue. Gibson and Walk tested 36 infants but only 27 actually moved, and Blakemore and Cooper reported detailed findings from only two kittens. With such small samples one anomalous baby or kitten could have a large effect on the pattern of results, so we cannot be sure the findings would be consistent in a larger group, even though the quantitative results themselves are clear.

Finally, much of the research extrapolates from animals. The chicks, lambs and kittens gave consistent results, but findings from animals may not reliably reflect human development because humans are born far less mobile and may rely more on learning. The fact that the human infants were all over six months old also means the research could not be repeated with newborns, so the reliability of the 'innate' conclusion cannot be checked in the same way. In conclusion, research into perceptual development is generally reliable because it uses standardised apparatus and quantitative measures that have been successfully replicated, but small samples and subjective observation of animals and babies mean some findings should be treated with caution.

5(c) A psychologist would first explain to Nina that perception develops through active exploration, so play that involves handling, moving and looking at real objects will help far more than screen time. One strategy is structured block and shape-sorter play to develop visual form and

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

Candidate name: *StemUp Model Solution*

size constancy. Nina's staff could set up a daily session where the 2 to 3 year olds post shapes through a shape sorter and the 4 to 5 year olds build towers with blocks of different sizes and then sort them by size and colour. This encourages children to recognise that a shape stays the same even when it is rotated and that a block is the same size whether it is near or far, which is the basis of the perceptual constancies. The psychologist would suggest staff use positive reinforcement, praising or giving a sticker each time a child posts the correct shape, because operant conditioning shows that behaviour that is rewarded is repeated, so the children will keep practising.

A second strategy is physical play that develops depth perception. Following Gibson and Walk, who found infants could perceive depth and avoid a visual cliff once they could crawl, Nina could use a soft-play area with low steps, slopes and a small climbing frame with a mat below, so the children practise judging drops and distances safely. Games such as throwing bean bags into hoops placed at different distances also train the children to use cues such as relative size and height in the visual field to judge depth.

A third strategy is music and listening games to develop auditory perception. Nina could run a circle time where the children listen to a song and clap the rhythm, or play 'what made that sound?' with hidden instruments, developing auditory constancy and the ability to pick out one sound from a background. Staff should model the behaviour first, because Bandura's social learning theory shows young children imitate adults they like, so if the staff enthusiastically join in the children will copy them.

These strategies are cheap, safe and suitable for 2 to 5 year olds, and because they are fun they give the children a real alternative to screen time while actively developing their perceptual skills.

6(a) Hall and Player aimed to find out whether being given an emotional context about a crime would affect a fingerprint expert's analysis of a print, and whether the experts felt they had been emotionally affected by the circumstances of the case. The sample was 70 volunteer

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

Candidate name: *StemUp Model Solution*

fingerprint experts from the Metropolitan Police Fingerprint Bureau, whose experience ranged from under three months to over thirty years, with a mean of eleven years. It was a field experiment with an independent measures design. A poor quality print of a volunteer's right forefinger was digitally superimposed onto a £50 note. Each expert received an envelope with the test mark, the ten-print fingerprint form of the 'suspect' and a crime scene examiner's report. In the low-context condition the report said the suspect had tried to pay with a forged £50 note, while in the high-context condition it said the suspect had fired two shots at the victim before running away. The experts worked in their normal office in pairs and had as long as they wanted. They recorded whether the mark was an identification, not an identification, insufficient detail, or some agreement but not enough to individualise, and then completed a feedback sheet asking whether they had read the report and whether it had affected them. The results show that bias affects the processing of forensic evidence much less than might be expected, at least for experienced professionals. Although 52% of the murder group said they felt the emotional context had affected their analysis, compared with only 6% of the forgery group, there was no significant difference in the actual decisions made: 20% of the low-context experts and 17% of the high-context experts were confident enough to present the mark in court as a positive identification, and most in both groups said there was insufficient detail. Thirteen experts (19%) did not even read the crime scene report, so for them context could not have had any effect. This suggests that experts believe they are being biased but their training allows them to focus on the physical evidence, so the perceived bias does not translate into a real effect. However, the study used a mark that most experts judged to have insufficient detail, so it may not show what happens with more ambiguous prints, where other research such as Dror et al has shown context does change expert decisions.

6(b) Science involves objective, empirical observation, controlled testing of falsifiable hypotheses, collection of quantitative data and replicable procedures so that cause and effect can be

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

Candidate name: *StemUp Model Solution*

established. Research into the collection and processing of forensic evidence is scientific to a reasonable extent, but there are also important limits.

Hall and Player's study has many scientific features. It was an experiment with a clearly manipulated independent variable (the forgery or murder context) and an independent measures design, with all experts receiving the same superimposed print, the same ten-print form and the same instructions, so the procedure was standardised and could be replicated. The dependent variables were operationalised as categorical decisions (identification, not an identification, insufficient detail) and the data were analysed with a chi-square test, so the conclusion that context did not significantly change decisions (17% versus 20% positive identifications) is based on objective, quantitative evidence rather than opinion. This makes the research falsifiable, because if context did bias experts the data would have shown it. However, the study was conducted in the field, in the experts' normal office with colleagues around them, which lowers control. The experts could take as long as they liked and worked in pairs, so extraneous variables such as discussion with a partner or distractions could have affected their decisions, making it harder to be certain about cause and effect. The measure of whether the context 'affected' them was also a self-report feedback sheet, which is subjective: the 52% who said they were affected may have been giving a socially desirable answer, and there is no objective way of checking an internal feeling. In this sense, part of the data is not empirical at all.

Dror et al's research on fingerprint experts is arguably more scientific in its controls. In one study five experts were given a print they had themselves matched years before, but were told it was the pair from the Madrid bombing misidentification; four of the five changed their original decision. This used a within-participants comparison against each expert's own earlier judgement, giving a highly controlled demonstration of contextual bias. Yet the sample of five is tiny and the deception means the procedure is hard to replicate widely, so the generalisability of this scientific finding is limited.

Other research in this area is much less scientific. Charlton, Fraser-Mackenzie and Dror

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

Candidate name: *StemUp Model Solution*

investigated motivating factors in fingerprint analysis using semi-structured interviews with thirteen experts and thematic analysis. The rich qualitative data gave insights into emotions such as the satisfaction of catching criminals and fear of making errors, but the themes depend on the researchers' subjective interpretation, cannot be falsified and would be difficult for another team to reproduce exactly. Miller's work on the line-up approach to hair comparison, by contrast, used an experimental comparison of standard versus line-up methods, giving measurable error rates, which shows that even applied forensic procedures can be tested scientifically.

A final point is that in forensic work the 'ground truth' is often unknown, so research can only measure consistency between experts rather than accuracy, which limits how far it can claim to establish objective fact.

In conclusion, the core research into forensic evidence, particularly Hall and Player, uses experimental designs, quantitative data and standardised procedures and is therefore largely scientific, but field settings, reliance on self-report and the use of qualitative interview studies mean the area as a whole is only partly scientific.

- 6(c) *A psychologist would explain to the experts that bias in forensic analysis is usually unconscious, so the strategies need to control the flow of information rather than rely on the experts simply trying harder.*
- The first strategy is Linear Sequential Unmasking, developed from Dror's research on contextual bias. The fingerprint experts should examine and document the features of the latent print recovered from the crime scene before they are shown the suspect's prints, and only then compare the two. Any changes they make to their initial analysis after seeing the suspect's print must be recorded and justified. This stops the suspect's print from guiding what the expert 'sees' in the crime scene mark, a form of confirmation bias, and it fits the ACE-V procedure (Analysis, Comparison, Evaluation, Verification) that the experts already use.*
- Second, the police should withhold the crime scene report and any emotional details of the*

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

Candidate name: *StemUp Model Solution*

case from the experts. Hall and Player found that over half of experts told they were dealing with a murder felt the context had affected them, so the experts should receive only the information needed to do the comparison, not what the suspect is alleged to have done or whether they have confessed. This can be managed by a case officer who filters the information and passes on only the evidence itself.

Third, the handwriting expert analysing the note should work with a line-up or filler-control method, as suggested by Miller for hair samples. Instead of being given the note and the one suspect's writing, the expert is given samples from the suspect and several fillers and asked which, if any, matches. If the expert picks a filler, this reveals the match is unreliable, and it removes the expectation that the single sample provided must be the culprit's.

Finally, verification should be blind. A second expert should check the fingerprint and handwriting conclusions without knowing what the first expert decided, especially if the first expert is more senior, since conformity to a senior colleague's view is a known source of bias. The experts could also receive training on their own biases, using examples such as Dror's study where four out of five experts changed their decision when given misleading context, so that they accept the need for these procedures.

Together these strategies reduce bias in both the collection and processing of the fingerprint and handwriting evidence while remaining ethical and practical for the police to apply.

- 7(a) Ulrich aimed to investigate whether having a view of a natural scene through a hospital window would have positive effects on patients' recovery from surgery, compared with a view of a brick wall. The study used secondary data from the records of 46 patients who had undergone a cholecystectomy (gall bladder removal) in a suburban Pennsylvania hospital between 1972 and 1981. The patients were in rooms on the second and third floors of a wing where some windows looked onto a small stand of deciduous trees and others onto a brown brick wall. The rooms were nearly identical in size, window size, bed arrangement and furniture, and the same nurses were assigned to each floor. Ulrich used a matched pairs

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

Candidate name: *StemUp Model Solution*

design: 23 patients with a tree view were each matched with a wall-view patient on sex, age within five years, smoking, obesity, general nature of previous hospitalisation, year of surgery and floor level. A nurse with extensive surgical experience, who did not know which view each patient had, extracted five types of data from the records: number of days in hospital, number and strength of analgesics each day, number and strength of doses for anxiety, minor complications such as headache and nausea, and all nurses' notes about the patient's condition.

The findings show that a window view may influence recovery to a considerable extent. Patients with the tree view spent less time in hospital after surgery, on average 7.96 days compared with 8.70 days for the wall group. The tree-view patients also needed fewer doses of moderate and strong painkillers between days two and five, with the wall group taking far more strong analgesics such as narcotics, while the tree group took more weak analgesics such as aspirin. The tree-view patients had slightly fewer minor post-operative complications and received far fewer negative comments in nurses' notes, 1.13 per patient compared with 3.96 for the wall group, with notes such as 'upset and crying' and 'needs much encouragement' being more common in the wall group. There was no significant difference in anti-anxiety medication, so the view did not affect every aspect of recovery, and the effect was found only in a single hospital with one type of surgery and a fairly dull wall view. Nevertheless, Ulrich concluded that a natural view reduces stress and therefore speeds recovery, so hospital design should give patients views of nature where possible.

7(b) Research is useful when it has practical applications that improve people's lives, when it influences policy or leads to further research, and when its methodology is strong enough for its findings to be trusted. Research into the psychological effects of the built environment is useful in many of these ways, but its usefulness is limited by some methodological issues. Ulrich's study of the view through a hospital window is highly useful because its findings have been applied directly. Patients with a view of trees left hospital almost a day earlier (7.96

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

Candidate name: *StemUp Model Solution*

compared with 8.70 days), needed fewer strong painkillers and received fewer negative nurses' comments. This has influenced evidence-based design in healthcare, with new hospitals placing beds near windows, creating healing gardens and using nature images where real views are impossible. Because hospital stays cost a great deal, even a small reduction saves money as well as suffering. The study is also useful because it stimulated a great deal of further research into restorative environments, such as Ulrich's later work showing that nature pictures lower stress in blood donors.

However, the usefulness of Ulrich's study is reduced by its methodology. It used secondary data from the records of only 46 patients in one suburban Pennsylvania hospital, all recovering from the same gall bladder operation, so the findings may not generalise to other patients, operations or cultures. The comparison was between trees and a brick wall, so it is not clear whether any interesting view, such as a busy street, would have had the same effect. The data on patients' mood came from nurses' notes rather than objective measures, so there could be bias in what nurses recorded. These limitations mean architects should apply the research with caution.

Newman's work on defensible space is another useful example. He argued that housing with clear territorial markers, natural surveillance and no 'unowned' spaces would reduce crime, and his comparison of the Van Dyke and Brownsville housing projects in New York supported this. The ideas have been applied in urban renewal and in Secured by Design guidelines in the UK, so the research has clear practical and policy value. Yet the Van Dyke and Brownsville comparison was a natural experiment with no control over variables such as residents' income, so the causal link is uncertain, which limits how confidently policy can be based on it.

Research into the effects of urban renewal on wellbeing, such as Halpern's work on Easterhouse in Glasgow, where improving the housing estate reduced residents' anxiety and depression, is useful because it shows that spending on the built environment can improve mental health. However, such studies rely on self-report questionnaires that may be affected by social desirability and are difficult to replicate because every estate is different, so they

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

Candidate name: *StemUp Model Solution*

provide guidance rather than firm rules.

There is also a question of social sensitivity. Findings that high-rise living or dense housing harms wellbeing can stigmatise the people living there, and they can be used to justify demolishing homes, so usefulness has to be balanced against the impact on communities.

In conclusion, research into the built environment is very useful because it has directly changed the way hospitals, housing and public spaces are designed, but its usefulness is limited by small samples, natural experiments with weak control and reliance on self-report, so it is most useful when combined with further research in different settings.

7(c) *A psychologist would advise Sam to design the apartments so that they reduce stress and give residents contact with nature, privacy and a sense of control, because research shows these features improve wellbeing.*

The first strategy is to give every apartment a view of nature. Ulrich found that surgery patients whose window looked onto trees recovered almost a day faster and needed fewer strong painkillers than patients facing a brick wall, so a natural view reduces stress. Sam should fit the apartments with large windows facing any trees or parks in the town centre, add balconies on each floor so residents can grow plants in pots, and turn the roof of the twelve-storey block into a communal garden with trees, seating and a lawn. Where a window faces another building, planting climbing plants on the wall opposite or placing nature murals in the corridors would give some of the same restorative effect.

The second strategy is good soundproofing. Research such as Glass and Singer on unpredictable noise and Cohen et al on children living under flight paths shows that uncontrollable noise raises stress, lowers concentration and reduces willingness to help others. Because the block is in the town centre, Sam should use double or triple glazing facing the road, insulate the walls and floors between apartments and place bedrooms on the quieter side of the building, so residents can control the noise they are exposed to and sleep well.

The third strategy is to apply Newman's defensible space principles so residents feel safe and

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

Candidate name: *StemUp Model Solution*

know their neighbours. Rather than one long corridor serving all the apartments on a floor, Sam should divide each floor into small clusters of three or four apartments sharing a semi-private lobby, so the space is clearly 'owned' by those residents and strangers are noticed. Entrances should be limited and overlooked by windows to allow natural surveillance, and there should be no unowned spaces such as dark stairwells. This reduces fear of crime and increases social contact, both of which improve wellbeing.

Finally, the psychologist would advise avoiding overcrowding and protecting privacy. The apartments should be generously sized with separate bedrooms rather than completely open-plan layouts, because Calhoun's work on density and research on crowding in high-rise housing show that lack of personal space increases stress and withdrawal from neighbours. Together these design features would make the apartments healthier places to live and would be attractive to the people choosing to move in.

- 8(a) *Lewis et al aimed to examine the effect of dance on mood in older adults, specifically in people with Parkinson's disease (PD) compared with a control group, over both a long cycle of a ten-week dance programme and a short cycle of a single class. The sample was self-selected and consisted of 37 participants aged 50 to 80 years: 22 with a diagnosis of PD (12 men and 10 women) and 15 controls (7 men and 8 women), many of whom were carers for the PD participants. The study was an experiment with a matched pairs design for the PD versus control comparison and a repeated measures design for the long versus short cycle, with mood scores as the dependent variable. In week 1 participants gave informed consent, completed a demographics questionnaire and the Mini Mental State Examination, and completed the Profile of Mood States (POMS) to give a baseline. In weeks 2 to 11 they attended a weekly 50-minute dance class based on rhythmic dancing to a strong beat, with the style changing every two weeks, for example Bollywood, tango and cheerleading. In week 9 participants completed the Brunel Mood Scale (BRUMS) immediately before and after the class, rating how they felt 'right now', to measure the short cycle. In week 12 the POMS was completed*

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

Candidate name: *StemUp Model Solution*

again a few days after the final class to measure the long cycle.

The findings suggest that dance can improve mental health to a considerable extent, as long as mood is accepted as an indicator of mental health. Over the long cycle, total mood disturbance on the POMS fell significantly for both the PD group and the control group, and anger scores in particular were reduced in both groups, so the ten-week programme improved mood whether or not participants had PD. In the short cycle a single class produced significant improvements in anger, tension and overall mood disturbance, and fatigue was reduced for those participants who had been high in depression at baseline. However, the depression subscale itself did not reach significance in the short cycle, so one class is not enough to lift depressed mood, and because there was no non-dancing control group it is possible that the social contact and music, rather than the dance itself, caused the improvement. Overall the study shows dance has a real but limited effect on mental health, improving mood and anger over time, which is why Lewis et al suggested dance as a low-cost therapy for people with PD.

- 8(b) Validity refers to whether research measures what it claims to measure and whether its conclusions are accurate. It includes internal validity (whether the IV really caused the change in the DV), ecological and population validity (whether findings generalise to other settings and people) and the validity of the measures used. Research into exercise and mental health has strengths in some of these areas but weaknesses in others.*
- A major issue is the validity of the measures. Lewis et al measured mental health using the POMS and BRUMS, both self-report mood questionnaires. Participants rated how they felt 'right now' before and after a dance class, so social desirability and demand characteristics are likely: participants who had chosen to join a dance programme and enjoyed the music may have reported feeling better because they expected to. Mood is also not the same as mental health, so a reduction in anger scores does not necessarily show that depression or anxiety disorders improved. On the other hand, the POMS is a standardised, validated scale that has*

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

Candidate name: *StemUp Model Solution*

been widely used in sport psychology, and self-report is the most direct way to measure feelings, so the measure has reasonable construct validity.

Internal validity is also limited. Lewis et al had no non-dancing control group, so the improvement in mood over ten weeks could have been caused by the social contact of the weekly group, the music, the attention from the instructor or simply the passage of time rather than by exercise. Because the control participants were often the carers of the PD participants they were not independent, and the number of classes each person attended was not controlled. This makes it hard to conclude that dance itself caused the improvement. However, the study did include the short-cycle before-and-after measure in week 9, which isolates the effect of a single class more clearly, and the use of two different time scales strengthens the conclusion.

Population validity is low. The sample was self-selected and consisted of only 37 people aged 50 to 80 from one area, most with PD. People who volunteer for a dance programme are likely to already enjoy exercise, so the findings may not apply to people who dislike exercise, to younger people or to people with other mental health conditions. The small sample also makes it difficult to detect effects, which may explain why the depression subscale was not significant in the short cycle.

Ecological validity is a strength of the research, because the dance classes were real weekly sessions in a community setting with a real instructor, so the findings reflect how dance would be used as a therapy. In contrast, Pretty et al's green exercise research, where participants ran on a treadmill in a laboratory while pictures of pleasant or unpleasant rural and urban scenes were projected on a wall, is more controlled and found that pleasant rural scenes improved self-esteem and mood more than exercise alone, but running in front of a screen is artificial and may not reflect the experience of exercising outdoors.

There is also a risk of researcher bias, since the researchers in Lewis et al knew which group each participant was in and were enthusiastic about dance as a therapy, and the interpretation of subscale results could have been influenced by this expectation.

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

Candidate name: *StemUp Model Solution*

In conclusion, research into exercise and mental health has reasonable ecological validity and uses validated measures, but the reliance on self-report, the absence of proper control groups and small self-selected samples limit its internal and population validity, so the claim that exercise improves mental health is plausible but not yet firmly established.

8(c) A psychologist would suggest to Amit that regular exercise reduces stress and improves mood, so a short daily programme in the weeks before the tests could help the Year 6 pupils feel calmer and more positive.

The first strategy is green exercise. Pretty et al found that exercise combined with exposure to pleasant natural scenes improved mood and self-esteem more than exercise alone, and Ulrich showed that contact with nature reduces stress. Because the school is near open countryside, Amit could take the class for a brisk 30-minute walk or run along a footpath into the countryside two or three times a week, perhaps at the start of the day so the children arrive in lessons alert and calm. He could add a purpose, such as a nature scavenger hunt or a conservation task in the school grounds like planting a hedge, so the children stay engaged and gain a sense of achievement. Walking in a group also gives the pupils time to talk, which provides social support before the tests.

The second strategy is a dance programme. Lewis et al found that a ten-week programme of weekly 50-minute classes of rhythmic dancing to a strong beat, changing style every two weeks, significantly reduced anger and total mood disturbance, and that even a single class improved tension and anger. Amit could run a 30-minute dance session in the hall twice a week in the half term before the tests, changing the style regularly, for example street dance, cheerleading and Bollywood, to keep the 10 and 11 year olds interested. Dancing to music is aerobic, which releases endorphins and reduces stress hormones such as cortisol, and it is enjoyable, so pupils who dislike competitive sport will still take part. Amit could also use a short dance or movement break on the morning of each test, because the short-cycle findings suggest an immediate lift in mood.

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

Candidate name: *StemUp Model Solution*

A third strategy is simple yoga or stretching and breathing sessions at the end of the day.

Yoga has been shown to reduce anxiety, and the breathing techniques give children a tool they can use in the exam room itself when they feel anxious.

Whichever strategy Amit uses, the psychologist would advise making it regular, fun and non-competitive, and obtaining parental consent and risk assessments for trips into the countryside, so that the exercise genuinely improves the mental health of the whole class rather than adding pressure.