### ANSWER ONE – Turning to Crime/Upbringing

<table>
<thead>
<tr>
<th>Usha - top band 8-10 marks</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) How can upbringing in a disrupted family explain criminal behaviour? [10]</td>
<td>This answer directly addresses the impact of upbringing in a disrupted family on criminality. The answer is detailed and accurate (e.g., description of disrupted and comparison samples) and well elaborated (explains the three different measures of criminality used by Farrington and Juby and also provides data to show that the effect of family ‘disruption’ depends on whether it is the father or the mother who has left home). The answer is well structured and highlights the researchers’ conclusions in relation to the research question. The key concepts of ‘disruption’ and ‘conflict’ could have been explicitly defined at the start, rather than leaving the reader to work out the meanings of the terms from the ways in which they were used. Nevertheless this answer is firmly in the top band.</td>
</tr>
</tbody>
</table>

Many psychologists have attempted different explanations of criminal behaviour. One of these explanations is upbringing in a disrupted family. Juby & Farrington conducted a study which aimed to disentangle the links between family disruption and delinquency. The participants and data which were used were from Farrington’s large longitudinal ‘Cambridge Study in Delinquent Development’, which followed up 411 London males from age 8 to age 48-50, in order to investigate risk factors in offending. The study by Juby & Farrington involved 75 White-British males born around 1953, who grew up in an inner city area of London and came from disrupted families and a comparison group of 344 White-British males born at the same time and who grew up in the same location, who came from intact families. The disrupted and the comparison groups were roughly similar in terms of age, ethnicity and social class (mainly working-class) and came from the same location, since all the boys were participants in Farrington’s large longitudinal study. The dependent variables measured were the number of adult convictions of the P’s, the number of juvenile convictions of the P’s & the number of self-reported juvenile delinquencies. Results of the study showed that when comparing the data of both groups (disrupted family & intact family) there was a higher percentage scored on each of the three measures of the dependent variable of criminality in the disrupted family group than the intact family group. Also, amongst the P’s in the disrupted families, higher percentages were scored on each measure of criminality if the P’s mother was absent, compared to when the mother was present. |
Additionally, in the intact families group, higher percentages were scored on each measure of criminality if there was high conflict in the P’s intact family, compared to intact families where there was not conflict. Looking at the results, Juby & Farrington concluded that criminal behaviour can be explained by disrupted families as when children grow up in disrupted families they are more likely to turn to crime- as shown by the findings. People are also more likely to turn to crime when the mother is absent from the family, and when there is high conflict in the family.

<table>
<thead>
<tr>
<th>TOP BAND</th>
<th>Total words = 348</th>
</tr>
</thead>
</table>
(b) Evaluate the use of longitudinal research when considering upbringing as an explanation of crime. [15]

Longitudinal research has been used in some studies to explain crime in terms of upbringing, and Juby & Farrington's study is an example. They used a cohort method in their study to look at a group of people for a long period of time, and there were a few advantages to this method. The use of this longitudinal method allowed the researchers to follow the P’s over a long period of time and follow up on them, so any change in behaviour can be seen and causes can be identified more easily - i.e. whether the behaviour is being caused by the independent variable or some other extraneous variable.

This advantage is not present for a cross-sectional snapshot study method, which was used in Akers’s study on investigating the effects of parental and peer role models influencing teenagers in the USA to use alcohol/drugs. The snapshot method used here only looked at the P’s (around 3500 teenagers in Mid-West USA) for a very short amount of time, so a cause-effect relationship between the IV and DV is harder to infer as an extraneous variable could have affected the P’s responses on the Likert scales regarding use/abstinence of drugs/alcohol. For example, the P’s may have recently watched a TV advert (before the study took place) which discouraged drinking - which could have affected the results. This means the longitudinal method is more useful in inferring causal relationships.

It would have been useful to define ‘cohort method’ at this point. A ‘cohort’ means in this context a group of individuals born at around the same time, who all belong to the same generation.

This is an important point, but really needs an example from the study to back it up. One of the more interesting findings from Farrington’s main longitudinal study (the ‘Cambridge Study of Delinquent Development’) was that the vast majority of boys who grew up on an East London Council Estate in the 1950’s/60’s engaged in some form of petty criminal behaviour and that criminality peaked between the ages of 17 and 25 years. By their mid-30’s, however, the majority of participants had desisted from crime, maybe because at this stage in their lives the ‘reward-cost’ balance had shifted in favour of behaving as a law-abiding citizen. This is an excellent example of a ‘cohort’ (generational) effect, which was highlighted by the longitudinal method of study.

This paragraph develops the first point, making a very effective contrast with the use of a snapshot approach to considering upbringing as an explanation for crime. This time the point is effectively illustrated by means of an example – the snapshot study by Akers which investigated predictions.
On the other hand, longitudinal studies are a very long process which can go on for many years before researchers can get results, analyse them and gather conclusions.

The snapshot study method however, is a quick way of looking at a proportional representation of the target population and results can be gathered quicker, and conclusions can be made. The fact that the snapshot study is faster means researchers don’t have to wait years to bear fruit for their research. This makes the snapshot approach more useful as actions can be taken based on conclusions from research, in order to help prevent crime.

Another disadvantage of longitudinal studies is that from social learning theory of a relationship between the use of drugs and alcohol by adult and peer role models and the use of these drugs by teenagers. (The use of drugs and alcohol might be considered ‘deviant’ behaviour which might lie on the ‘slippery slope’ towards criminal activity. Ideally, the candidate should have spelt out this assumption, rather than leaving the reader to wonder about the precise relevance of this example to the topic of criminality.)

The candidate now turns to a limitation of the longitudinal approach. She argues very effectively that the long time-scale of longitudinal research potentially makes such studies less useful than snapshot studies in providing immediate conclusions which can be used to design social policies to reduce crime. Once again, an opportunity to discuss an interesting example has been lost.

In fact, the 40-year Farrington longitudinal study is still providing evidence for social policy interventions to help youngsters who are ‘at risk’ of turning to crime, for example pre-school intellectual enrichment programmes, family therapy and parenting skills and mentoring programmes for adolescents.

Farrington (2011*) has highlighted the influence of his own research in shifting government priorities away from physical/situational crime prevention (1980’s) towards early intervention and ‘nipping crime in the bud’.
there are often significant drop out rates, as seen in the Farrington cohort study, where many of the P's had dropped out of the study, so results were based on a smaller number of P's than from the beginning. If there is a high drop-out rate, this reduces the 'population validity' (representativeness) of the results from this sample. If the sample size becomes smaller, then results may be biased by anomalies in the sample, so it would be wrong to assume that findings which were true of the sample were applicable to everyone else. On the other hand, snapshot studies don't have this problem because the method is quick so the study can be completed before anyone decides to drop out (which usually happens in longitudinal studies midway through the study- i.e. a few months/years in - just like in the Farrington study).

In conclusion, the use of longitudinal research is useful in explaining crime in terms of upbringing because causal relationships can be inferred more easily without the worry of extraneous variables, but snapshot study methods are more useful when looking at large samples, as the drop-out rates aren't as high, since the snapshot method makes the study quick.

Usha has made a valid general point, but she has made a factual error in claiming that there was a 'significant' drop-out rate in the Farrington study. In fact this study was remarkable for its high retention rate – families seemed to take a pride in being part of the study.

Beware of using a formulaic check-list for evaluation points, regardless of whether they actually apply to a given study!

ANSWER TWO  Making a Case /Interviewing suspects

<table>
<thead>
<tr>
<th>Yasmin TOP BAND 9-10 marks</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Describe one piece of research into how lies can be detected</td>
<td>This is a very clear summary. The level of detail and elaboration on sample and procedure is excellent. Interesting details are provided on the conclusions drawn by the researchers.</td>
</tr>
</tbody>
</table>

Research carried out by Mann et al. in 2004 aimed to test police officers' ability to distinguish between a truth and a lie when interviewing a suspect. They carried out a field experiment where 99 Kent police officers (24 female, 75 male) were asked to judge the truthfulness of suspects in real-life police interviews. The sample of police officers included 78 detectives, 8 trainers, 4 traffic officers and 9 uniformed response officers. The police officers had to fill out a questionnaire about their experience in detecting lies. They were then shown video clips of 14 suspects, showing the head and torso, so that ‘body-language’ such as facial expression and movement was visible. The 54 clips used were selected so that each clip was known from other evidence to be either entirely truthful or entirely a lie. There were 23 truthful clips and 31 deceptive clips. Clips varied in length from 6 to 145 seconds. After watching the clips the police officers had to state whether they thought the suspect was being honest or not and then state how confident they were in their answer. They were also asked the cues they used to detect lies. The results from the study revealed that police officers performed better than chance (50%) as they were 66.2% accurate on lies and 63.6% accurate on truths. The most frequently mentioned cues were gaze, movement, vagueness, contradictions in stories and fidgeting. This meant that the researchers were able to conclude that police officers do have an ability to detect lies; this ability got better as they became more experienced. They also concluded that the police officers who were better at detecting lies used story cues more than body language cues. |
### b) Discuss the validity of information gained from suspects in interviews [15]

Information gained from suspects in interviews is extremely high in ecological validity, meaning that it really measures how people behave in everyday life or real life situations. For example, Mann’s study, which tested police officers’ ability to distinguish truth and lies during with interview with suspects was a field experiment as the police officers were real and were judging the truthfulness of people in real-life police interviews of real criminals. As they were real suspects, they would exhibit the cues that police officers would expect to see in their own interviews, allowing the police officers to apply the techniques they usually used to detect lies. In Gudjohnsson’s study, he documented a real case of false confession of a youth who was at the time distressed and susceptible to interrogative pressure. As these studies tested the skills of police officers in close to real life situations, this increases their ecological validity, enhancing the applicability of the conclusions to real life.

Arguably, information gained from suspects in interviews can be seen as being low in internal validity, which relates to how much control is used over extraneous variables. For example, in Mann’s study, the results showed that police officers performed better than chance, as they were 66.2% accurate on lies. However, no control group of lay people was used to compare to the police officers’ scores to find out whether their ability at detecting lies was better than the general public. A control group of lay people could not have been used, as they would not have been allowed to view the footage of real police interviews. This reduces the validity of Mann et al.’s results because they could not compare the ability of the police interviewers against a control group. The aim was to see if police officers were any better than others at detecting lies, but this cannot be achieved without a control group.

Including a control group in this research would break ethical...
Guidelines on protecting the identity of the people in real-life interviews. This represents a contradiction between being experimentally valid and being ethical. In these circumstances it is more important not to break ethical guidelines on the protection of participants from psychological harm than to conduct experimentally valid research.

Acceptable practice. She understands that there may be good reasons to conduct research which is less than ideal in design, and makes the judgment that ethical practice come first and foremost.

Can you suggest any way to include a control group in this study but at the same time adhere to ethical guidelines? Would this have any drawbacks?

The measurement of cognitive processes in the information gained from suspects in interviews can also negatively impact the validity, as cognitive processes can't be seen and measuring how someone thinks or processes information is a difficult task.

For example, in Mann's study, the police officers had to detect whether the person in the real-life interview was lying or telling the truth, which is a cognitive process. This was hard to measure and was measured through "cues", such as gaze, fidgeting, contradictions in stories and vagueness, which could be subjective, depending on the policeman's experience or viewpoint.

The "cues" observed may also actually measure a person's level of nervousness and not detect whether they are lying or not. Therefore it may be suggested that the validity of measurement of the cognitive processes may be affected through mis-reading "cues".

In this third paragraph, Yasmin looks at the validity of the ways in which psychologists attempt to assess the 'cognitive process' of lying.

She is implicitly defining validity as 'whether we succeed in measuring what we intend to measure'.

She explains very clearly why 'body language' might merely indicate stress levels and therefore not be a valid measure of truth-telling or lying.

In conclusion, the validity of information gained from suspects in interviews is quite mixed, as research can be extremely high in ecological validity by being conducted in real life situations, but can also suffer from low internal validity, through the lack of controls (in Mann's study). The measurement of cognitive processes in tests on the information gained from suspects in interviews can also negatively affect the validity, as cognitive processes can't be seen and are therefore hard to measure.

This conclusion effectively summarises the argument of the three preceding paragraphs.

Total words = 593
This answer was not written under exam conditions, so it is rather longer than could be realistically achieved in 20 minutes.
This answer focuses on the validity of the research methods used to investigate how suspects are interviewed. The answer effectively combines breadth with depth. Yasmin explores three different aspects of validity, using research evidence very effectively to substantiate the points she makes. The answer shows conceptual clarity and develops an argument. She debates the important issue of whether ethical principles should be sacrificed in the interests of validity. The concluding paragraph effectively summarises the points she makes.

The following two paragraphs are drawn from an alternative answer to the same question by another student from the same class as Yasmin. These paragraphs attempt a direct discussion of the validity of information from suspects, rather than the validity of the research methods used. This is harder to accomplish. The discussion is interesting, but rather less well focused than Yasmin’s answer.

A strength, in terms of validity of the information gained from suspects in interviews is that case studies on the issue provide lots of in-depth information about it and this has useful applications. For example, Gudjohnsson et al. found that a young man despite being innocent allowed himself to be coerced into confessing that he murdered two women, probably in order to escape the pressures of the interview. As this is a real life example, it is high in ecological validity; Gudjohnsson et al. were able to gather lots of qualitative data on what happened during his interview. Case studies are so useful because they are example of where it has happened in real life, therefore they could analyse what led to him allowing himself to be coerced into confessing. It is likely that through analysing examples like these Inbau came up with his ‘Nine steps in interrogation’.

A weakness in terms of validity of the information gained from suspects in interviews is that it can break ethical guidelines and cause the suspect psychological harm just to gain a confession. This is shown in Inbau et al.’s ‘Nine steps of interrogation’ where the police officer switches between directly accusing the suspect of the crime and then pretending to help them, by giving them a false dichotomy where both lead to their confessions of the crime. This may be effective at getting suspects who committed the crime to confess; however it can also lead to false confessions, because the suspect becomes convinced that they are guilty, due to the psychological stress they are placed under. An example of where this has occurred is in the study by Gudjohnsson et al. where they found that a young man despite being innocent, having average intelligence and no mental health problems allowed himself to be coerced into confessing that he murdered two women, in order to escape the gruelling interview. This shows that the approach that the interviewer takes when interviewing the witness can greatly affect the answer they receive.
### Answer THREE. Psychology of the Courtroom / Persuading the Jury

<table>
<thead>
<tr>
<th>Zamira - top band 9-10 marks</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a) What is the effect on a jury of evidence being ruled inadmissible in court? [10]</strong></td>
<td>A clear and concise account. The implication of this research could have been more clearly highlighted - that in some circumstances drawing the attention of a jury to inadmissible evidence by the use of a legal explanation may have the opposite effect of that intended, by apparently making it more difficult for jurors to ignore the inadmissible evidence.</td>
</tr>
</tbody>
</table>

The effect on a jury of evidence being ruled inadmissible in court has been studied by Pickel (1995). In Pickel's controlled laboratory experiment, he aimed to find out the effect of evidence being ruled inadmissible. Participants viewed a re-enacted court trial and inadmissible evidence was included 'accidentally' by the researcher. This critical evidence included information about past convictions. This was an independent measures design, as participants were either put into a group where the evidence was ruled admissible, inadmissible, inadmissible with a legal explanation, or had not heard any critical evidence by the judge. In the condition where the evidence was ruled inadmissible with a legal explanation, participants were told why they should ignore the information, for example 'it could be incorrect'.

In this study there were 528 participants, all from Ball University in the USA, who had participated for credit as part of their psychology course. Pickel found that in the condition where the critical evidence had been ruled admissible there were 42% guilty verdicts and where evidence was ruled inadmissible without a legal explanation there were 43% guilty verdicts. However, those participants who were told that the evidence was inadmissible and were given a legal explanation afterwards showed a much higher rate of guilty verdicts at over 60%. Pickel concluded from this that when attention was brought to the evidence that was accidentally leaked, the participants had clearly not been able to ignore it; and thus paid it more attention.

**TOP BAND**

Total words = 243
Explanatory note
In fact the study by Pickel (1995) comprised three separate experiments, of which Zamira has described only the first. Experiment 1 and Experiment 2 investigated the responses of student mock jurors to different types of inadmissible evidence. The story line for Experiment 1 included information about prior convictions of the defendant, whereas Experiment 2 included hearsay evidence (where the witness has heard someone else but has not personally witnessed those events.) The results for Experiment 2 were different to Experiment 1; participants in Experiment 2 seemed to be able to ignore hearsay evidence, whether they received a legal explanation or not. Pickel found that participants in both inadmissible conditions were less likely to find the defendant guilty than those in the admissible condition but did not differ from a control group (no ‘critical’ evidence) in terms of likelihood of finding the defendant guilty. So in this experiment, the legal explanation did not ‘backfire’ on the participants, as it had done in Experiment 1. Pickel suggests that participants already suspected that hearsay evidence should not be used, so that it was easy for them to disregard it when so instructed.

What are the practical implications of these apparently contradictory results? Is it advisable or not for judges who are obliged to rule courtroom evidence inadmissible to add a legal explanation? The answer seems to be that it depends on the type of inadmissible evidence. So although these results are very interesting and thought-provoking, it is impossible to use the results to make simple recommendations which apply in all situations. Also, this research was done within the American legal system, where rules and practices concerning inadmissible evidence are somewhat different to common practice in the UK, so caution is needed in applying these findings in the UK.
<table>
<thead>
<tr>
<th>Zamira - Top Band 12-15 marks</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(b) Evaluate the usefulness of research into persuading a jury</strong></td>
<td></td>
</tr>
<tr>
<td>Psychologists who do research into persuading a jury all aim to conduct experiments that are ecologically valid and also to use samples which representative of the wider population. However due to ethical reasons, this is not always possible.</td>
<td>Zamira makes a good start by referring at the outset to three different factors which could affect the usefulness of research into persuading a jury. However, it would have been even better to refer explicitly to ‘usefulness’ in her opening statement. Zamira knows what she is about to say, but she leaves the reader a little uncertain.</td>
</tr>
<tr>
<td>Research into persuading a jury commonly lack usefulness, due to the types of participants they experiment on.</td>
<td>Now she introduces the key word ‘USEFULNESS’ but she once again does not quite hit the target. She would have done better to EXPLICITLY state the GENERAL evaluation to be discussed (sampling and generalization/population validity) in this signpost paragraph. Zamira has started with her example, rather than the point it exemplifies.</td>
</tr>
<tr>
<td>They often consist of students, as are opportunity sample. In the USA it is a much known fact that university students are asked to participate in experiments in order to gain course credit. For example Pickel’s study consisted of all undergraduate students from Bali University in the USA. In addition, Pennington and Hastie’s study of the way the prosecution story is presented also consisted of a student sample; of 130 undergraduates from NorthWestern and Chicago University.</td>
<td>Point about limitations of using psychology students as SAMPLE, is aptly illustrated with two clear and explicit examples from different research studies.</td>
</tr>
<tr>
<td>Although students are far cheaper to get to participate in a study and easier, they reduce the usefulness as they lack ecological validity.</td>
<td>Did she mean ‘population validity’? No matter, because she has placed her point so clearly in context that we understand</td>
</tr>
</tbody>
</table>
This means that there is the problem of being able to generalize the findings of the study to the wider generation. This is because students may differ to the adults in the way they think, as students are all similar ages, class and attitudes.

Furthermore, students who are taking part in a study for course credit in their Psychology courses, ay show demand characteristics. As they are familiar with psychology they may figure out the aims of the researchers, and represent themselves in a way that will make them look good and of what is expected of them.

This is a disadvantage of having a student sample, and it therefore reduces validity as well as the usefulness of the study.

A further problem with research into persuading a jury is that they are usually carried out in laboratory conditions, whereby participants do not see a real trial as it would happen in a real case. For example in Pennington and Hastie’s study of story order and witness order, the 130 student participants listed to an audio tape of a re-enacted trial. This is a major problem of research into persuading a jury as it lacks mundane realism, meaning that the task is not similar to what it would be like in an actual court case. A jury is not asked to listen to an audiotape of the court case, but instead they watch it happen in front of them. This reduces the usefulness of the study as participants may realise that they are not part of a real jury and decide that they will not put a lot of effort into it as they are not sending down an actual criminal. Nonetheless, it is impossible to study the actual jury in a case due to ethical reasons as the jury
swears to secrecy that no information will be passed outside that room.

Therefore we can conclude that some research into the persuading of a jury is better than none and any information to juries and judges is useful.

For example, from Pennington and Hastie’s study, it would be useful to advise the prosecution to present the story in story order as this produced the highest number of guilty verdicts and the defendant in witness order as this produced the least guilty verdicts.

Effective comment on WHY mock trials may lack EV context of example
New point about ETHICS of conducting ecologically valid studies

Conclusion to part - with important comment - some useful research is better than none (Note how she presents a counter-argument to the view that the usefulness of all the research is compromised by methodological limitations)

Perhaps not the best positioning of a final example of the way in which research evidence on courtroom persuasion can be useful. Maybe this example should have been used as part of the opening paragraph rather than the final one?
Tony- lower band 3-5 marks

a) Describe ‘anger management’ as a treatment programme for offenders. [10]

Anger management is a treatment programme for prisoners. Anger management is a cognitive method of changing a prisoner's personality to make them less angry.

One study that shows the effect of anger management is a study by Ireland. Ireland collected 238 prisoners to undertake an anger management course called 'CALM'. The variable was tested through surveys given to the prisoners after the course, and prison officers also recorded and reported the behaviour of the prisoners, using Wing-based measures. Anger management was said to be useful from this study as results show that officers reported less angry incidents and prisoners also stated on their survey that they felt less angry and more relaxed.

Comment

The question calls for a description of anger management as a treatment programme. The question does not ask for a description of a research study about the effectiveness of anger management, so the description of the study by Ireland does not attract credit. This answer only achieves a mark in the lower band because only the first sentence directly answers the question.

This question is an example of a trap for the unwary. It is not correct to assume that the Forensic option requires knowledge of 36 research studies. What is required is knowledge of 36 pieces of psychological evidence, which may take the form of theories, or research studies, or as here, therapeutic techniques.

LOWER BAND 3-5 marks

Total words = 112
<table>
<thead>
<tr>
<th>Jay - Top Band 9-10 marks</th>
<th>a) Describe ‘anger management’ as a treatment programme for offenders. [10]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) ‘Anger Management’ is a cognitive based treatment programme that aims to reconstruct the faulty schemata (thinking patterns) that lead to criminal behaviour. Anger management is a well-established treatment programme for offenders and, with the rates of recidivism resting stubbornly at 64%, it has been set up in 24 prison establishments in England and Wales. One example is the programme ‘CALM’. This programme is suitable for offenders who have ‘hostile aggression’ and have no mental illnesses and it is necessary that they can read and write. By ‘hostile’ aggression I mean they react very explosively to a stimulus or object and are very impulsive. The programme aims to reach them techniques to manage their anger. Based on the ‘Novaco Inventory’, developed by Arnold Goldstein, there are three particular stages: Firstly, the must accept what it is that is making them resort to anger and learn to recognise the situations it occurs in. Secondly they must undertake relaxation programmes—breathing—and learn to displace their anger onto something else. Trainers will also model ‘cognitive alternatives’, eg walking away from the situation rather than just hitting out. Finally, they must learn to ‘role play’ these new techniques so they fully learn them and they become second nature. More specifically, anger management programmes offenders to identify the psychological effects of anger eg sweating, increased heart rate and aim to provide arousal management. Specific techniques to</td>
<td></td>
</tr>
<tr>
<td>This is an excellent answer which more than fulfils the criteria for the top mark band. The answer makes correct and comprehensive use of psychological terminology, such as ‘cognitive based treatment programme’, ‘schemata’, ‘hostile aggression’, and ‘cognitive alternatives’ and also provides examples of what is meant by these terms. A full description and explanation of three stages in anger management training is provided, as well as a clear and informative explanation of some more specific techniques. Full marks could have been achieved without the inclusion of the evaluation study by Ireland.</td>
<td></td>
</tr>
</tbody>
</table>
achieve this include 1) 'Thought stopping'—stopping negative thoughts in an 'interpersonal confrontation'. 2) 'Fogging'—hearing criticism but not reacting (in an interpersonal situation) and 3) 'Broken record' technique—calmly repeating what you want.

Ireland did an evaluation of the treatment programme 'CALM', using two groups matched on response in a cognitive interview and Wing behavioural check.

Anger Management programmes usually work for those with 'hostile' aggression, where the aggression is uncontrollable, rather than 'instrumental', where it is a means to an end. Both of Ireland's groups (experimental and control) were defined as having 'hostile' aggression. Ireland found that CALM was effective, and there was a significant difference, both in reported anger and in Wing behavioural checks, between the experimental group and control.

| TOP BAND 9-10 marks | Total words = 357 |
Jay - Upper Band  8-11 marks

b) Assess the effectiveness of offender treatment programmes [15]

Assessing the effectiveness of offender treatment programmes can be difficult, as there is a debate as to what exactly they should do. For example the Wheatley ear acupuncture investigation assessed the effectiveness of ear acupuncture for drug addicts. Wheatley compared the group who received the treatment with a control. He used quantitative and qualitative data to test effectiveness. Wheatley found that the offenders reported increased relaxation and better sleep. Some would argue that increased 'relaxation' does not show the effectiveness of the offender treatment programme, as the aim of prison is to punish. On the other hand, Wheatley did also find a 70% reduction in drug related incidents, suggesting that ear acupuncture is effective in prevention of drug related incidents. This brings me on to my next point.

'Proxy measures' for the effectiveness of therapies can be untrustworthy, as they are often susceptible to demand characteristics. For example, Ireland demonstrated the effectiveness of the CALM treatment programme through the use of behavioural checks and self reported aggression (proxy measures). These measures are useful, as they show how the offender regards the treatment; however the offenders may have been behaving more 'reformed' when behaviour was checked within the prison, in order to seek early parole.

Offender treatment programmes do have many strengths, as they can help change beliefs (e.g. 'CALM') as well as behaviour (Wheatley). However there are difficulties in measuring their effectiveness, due to demand characteristics.

This is a very interesting answer, which explores in some depth the important question of how to choose the most appropriate measure of the 'effectiveness' of treatment programmes, making apt use of examples from two research studies.

The writer is critical of the appropriateness of 'proxy measures' of treatment effectiveness. A 'proxy measure' is a substitute measure used under circumstances when it is difficult to measure directly. In the context of offender treatment programmes, the writer seems to be assuming (but does not make explicit) that a direct measure of effectiveness would involve monitoring criminal activity after the offender had left prison; therefore all measures taken within the prison setting (either self-report or behavioural measures of aggressiveness, or even more indirect measures of rehabilitation, such as improved relaxation after acupuncture) must be considered to be 'proxy' measures.

The answer is of potentially very high quality, but fails to attract a top band mark because it basically only explores one issue rather than a range of issues and also because of the failure to explicitly define the term 'proxy measure'.

UPPER BAND – 8-11 marks

Total words =240