November 2015 subject reports

Psychology

Overall grade boundaries

Higher level

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 8	9 - 19	20 - 31	32 - 43	44 - 55	56 - 67	68 - 100

Standard level

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 9	10 - 21	22 - 35	36 - 47	48 - 57	58 - 68	69 - 100

Higher level internal assessment

Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 2	3 - 5	6 - 9	10 - 13	14 - 17	18 - 21	22 - 28

The range and suitability of the work submitted

Overall the range and suitability of the work submitted was of good standard, with the majority of centres showing good understanding of the requirements of the internal assessment (IA). The majority of IA reports submitted met the criteria for an experimental design, whereby there was manipulation of an independent variable with the effects of a dependent variable recorded. Most of the more successful IAs were based on studies from the cognitive level of analysis, such as, effects of interference and memory recall, reconstructive memory, experiments related

to schema theory and imagery versus rehearsal. Some newer research was presented which was pleasing to see.

In general, the weaker reports shared the following characteristics:

- Background studies and/or theories were not clearly explained and/or made relevant to the hypotheses.
- The hypotheses were not operationalized, that is, made measureable.
- The use of the descriptive statistics was not explained.
- The target population was not clearly identified.
- The IV and DV were not clearly operationalized and made measurable.
- Discussions were superficial with little or no consideration of their results in the light of background research and/or no reference to statistics.
- Referencing was not a standard format or not complete.

It should be noted that it is not required to make an exact replication of an experiment. A partial replication is adequate. However, the candidate's experiment should be closely linked to a published experiment. A few candidates also 'combined' studies and replicated various aspects of each; this is strongly discouraged.

Candidate performance against each criterion

Criterion A: introduction

In many reports, the research presented was relevant and adequate in number, but could have been explained in more detail. Many candidates found it challenging to explicitly state how the research was highly relevant to the hypotheses. Candidates should also be aware that the background research should always logically lead towards the candidate's own research hypotheses which in turn allows for the formulation of a clear research hypothesis.

It is also critical that the variables stated in the hypotheses are operationalized (that is, made measurable) and a statement of significance made.

Criterion B: design

The design (repeated measures or independent groups) was not always properly justified. When identifying the IV and DV, candidates often did not operationalize them, that is, clearly make them measureable.

Overall, candidates had a good understanding of the ethical guidelines. It is permissible to use participants under the age of 16 if parental consent is given. This should be stated in this section when discussing ethical considerations.

Criterion C: participants

Candidates to neglected to state the target population, that is, the population from which the sample was drawn. Candidates also often confused the actual sample with the target population.

Most candidates did identify the appropriate sampling technique, but some struggled to explain the use of this method, for example, how the method was used.

Criterion D: procedure

It is necessary that all materials (for example, informed consent form, standardized instructions, etc.) are referenced in the appendices. Without proper referencing, it would not be possible to properly replicate the experiment. Candidates must also make clear how the control and experimental groups differed.

Criterion E: results – descriptive

Only the strongest candidates explained *the use of descriptive statistics*, that is, why the mean was chosen as the measure of central tendency and/or the standard deviation as the measure of dispersion. Most candidates included a graph and a table, but proper labelling of the graph was an issue with incorrect labelling or no labelling at all. Candidates must also present the results in narrative form as well as in a table and graph. Only one measure of central tendency and one measure of dispersion should be presented.

Criterion F: results – inferential

Many candidates chose an appropriate test and justified the use of the test (based on the level of data and the design). At times, t-tests were chosen (which is acceptable) but often it was not the most appropriate test based on the particular aspects of the experiment (sample and/or variance of data). The statement of significance should also always be appropriate and clearly stated.

It is important that all raw data and calculations of the inferential test are included in the appendices. If the calculation is performed online, a screen shot of the calculation should be included in the appendices as documentation.

Criterion G: discussion

Candidates should always refer back to *all* research and/or theories presented in the introduction and discuss these in reference to their own findings. Candidates who included research and/or theories in the introduction that were not highly relevant often struggled with this aspect of the discussion.

Almost all candidates presented limitations, but often in a superficial manner, without rigorous analysis. Limitations should be presented that are relevant to this particular investigation, not limitations of experiments in general. Also, certain aspects of the study that were presented as limitations could have been easily avoided with the use of a pilot study, thus indicating a limited amount of planning.

Modifications should be based on the relevant limitations of the study. It is also necessary that a conclusion is included for all reports.

Criterion H: citation of sources

Referencing continues to be an area of weakness. Candidates often did not cite all research or use a standard citation method (such as APA).

Criterion I: report format

Generally the report formats were well done. Appendices were well organized and labelled.

The abstract must include a summary of the study as well as the results of the study. This was often not included.

Recommendations for the teaching of future candidates

- Assist candidates in selecting an appropriate experiment to replicate with an appropriate theoretical framework and background research. Finding relatively simple experiments to replicate is recommended. Again, it is advised that candidates do a partial replication of studies rather than try to 'create' their own study. Only two variables should be manipulated.
- For the sample, the number of participants in the experiment does not need to exceed 20 (independent design) or 10 (repeated measures design), and it is recommended to observe this.
- It would be helpful if candidates were given published research to read in order to familiarize themselves with the aspects of experimental research.
- It is recommended that candidates be given guidance in accessing appropriate psychological journals. Many candidates only used internet sources of a non-specialist nature as background literature.
- Candidates and teachers should be fully aware of the assessment criteria to ensure that all guidelines are met.

Standard level internal assessment

Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 2	3 - 5	6 - 8	9 - 10	11 - 12	13 - 14	15 - 20

The range and suitability of the work submitted

The work submitted tended to vary in terms of quality and variety. Studies coming from the cognitive level of analysis were the most popular choice. Most of the reports submitted were a replication of studies on perception, cognition and memory topics in psychology. It was pleasing to note a trend towards replicating more recent research. In the majority of cases the work submitted was suitable for Diploma Programme candidates studying psychology at standard level and conducted within permitted ethical guidelines.

The majority of candidates selected appropriate studies, described them well and were able to somewhat link their own results to the findings of the original study in the discussion section.

Candidates often managed to score some marks in design and participant sections but sometimes did not include justifications and therefore could not be awarded full marks. In procedure and results sections some marks were not awarded due to lack of relevant and precise information.

In most reports many marks were lost in the discussion section due to lack of depth in discussing the findings and methodological issues of the conducted study.

There were some examples of reports that did not meet the criteria for experimental work, but these were very few.

Candidate performance against each criterion

There were some very solid samples showing a high level of ability. At the lower end, it was apparent that some candidates were appropriately instructed but failed in assigning enough time and energy for putting together an appropriate internal assessment (IA) report.

Referencing proved to be difficult for some candidates. Teachers should specifically instruct candidates that referencing should occur whenever a study/theory is described in the introduction.

Criterion A: introduction

In most reports, introductions were well written with most candidates clearly identifying and explaining the study for partial replication as well as presenting a clearly stated aim. However,

some candidates attempted to describe more than was necessary. Although a simple description of the study being replicated was all that was needed, too often candidates included superfluous material and studies, subsequently failing to clearly identify the study they were replicating. There is no need to provide a review of several studies in the introduction section.

Another problem that was occasionally encountered was that the introductions were often written in concordance with the higher level requirement of a literature review with the inclusion of hypotheses, thereby reducing available word count for the detail needed to describe the original experiment or making it more challenging to write a good and thorough discussion with a rather limited word count.

Criterion B: design

Although the statement of IV and DV was generally correct, including operationalization of both, the proper identification of the design itself was still problematic for candidates from many centres. Many candidates vaguely identified the design as just "experimental or laboratory controlled". It seemed that some candidates could not distinguish the design from the method. Also, a number of candidates provided incorrect justifications or no justification for their choice of experimental design.

In some reports the description of the IV and DV needs to be more clearly stated, they are often too vague and imprecise. Another rather common mistake was stating only one condition (usually the experimental condition).

There was a noticeable improvement in the identification and discussion of ethical considerations (informed consent, debriefing, etc.).

Criterion C: participants

In many cases candidates presented a good description including appropriate target population characteristics and identifying their sampling technique. However, many candidates did not justify the use of this sampling technique and therefore could not obtain full marks. The term "random" still tends to be a source of confusion reflected in the description of participant selection and allocation to conditions.

Criterion D: procedure

In the majority of cases procedures were relevant and clearly described, but in some cases materials referred to were not included in appendices (for example, standardized instructions, tests, questionnaires) which affected the replicability of the procedure. Although this section of the report was usually well done there is still some room for improvement – complete and detailed debriefing was rarely present.

Criterion E: results

Unfortunately, the results section was a weakness within many reports. In many instances graphs were not labelled clearly enough for conditions to be recognized. Weaker candidates chose the wrong type of graph (histograms or pie charts to show differences between independent groups). In addition, a number of candidates presented their results in an unclear

manner – they did not include percentages, measures of central tendency or dispersion. Some candidates provided several graphs in the results section - presenting the data in a variety of ways, but often not reflecting the aim of their study. Also, occasionally there was incorrect application of statistics. For example, when ordinal levels of measurement were used, there were some candidates who calculated the mean score in spite of the fact that this is not an appropriate measure of central tendency for ordinal data.

In the results section, candidates should ensure they provide table and figure headings and provide sufficient description of what these reflect. It is important that candidates specifically name their measures of central tendency; do these reflect mean, median, mode? Also, candidates should describe what these different scores for experimental and control groups reflect; and importantly what the standard deviation or range imply.

Some candidates made the mistake of graphing raw data. Another common problem was that candidates did not fully interpret their descriptive statistics. Calculations (for example, of mean) were sometimes inaccurate.

Criterion F: discussion

The quality of the discussion sections tended to vary. In this session there was some improvement as discussions tended to more clearly follow the criteria for this section and many more candidates linked the discussion of weaknesses to the type of design chosen. Conclusions tended to be embedded within the discussion section instead of just added on at the very end.

Unfortunately, those candidates who had not clearly described the study being replicated in the introduction tended to have difficulty with the discussion section as well.

Some reports failed to achieve higher marks because limitations were not clearly addressed and suggestions for further research were often omitted.

Criterion G: presentation

In general, reports were within the word limit (although a few times candidates hadn't recorded the word count). In the majority of cases reports used the required format and references were provided. However, full publication details of the replicated study were often not given in a consistent or full manner. Candidates should be encouraged to adhere to one standard referencing system. At times it seemed that some candidates finalized their reports in a hurry and therefore some items were omitted from appendices (for example, materials used, standardized instructions, and consent form).

Recommendations for the teaching of future candidates

- Teachers must be clear on what the basic requirements of the IA are in regard to which topics/experiments are not appropriate for replication due to ethics so that they can guide candidates to make more appropriate choices.
- Choosing an experiment within a basic level of knowledge and focusing the aim to what is manageable for the assessment criteria is of most importance.

- More instruction on the advantages and disadvantages of using various experimental designs and sampling techniques will help candidates justify the use of them in their reports.
- An experimental study for SL should not have more than one independent variable and one dependent variable.
- The results section should clearly provide descriptive statistics related to the aim of the study.
- Candidates should be encouraged to check all calculations and include them in the appendix.
- More emphasis should be put on the importance of a well-balanced discussion that makes explicit connections between the methodology and the results of their study. Candidates must have a balanced explanation of what they felt were strengths and weaknesses.
- More guidance is necessary in relation to the expected format for the internal assessment (for example, knowing where ethical considerations should be addressed, raw data presented, where standardized instructions belong).
- Candidates should be encouraged to proofread their reports before handing them in.

Higher level paper one

Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 3	4 - 7	8 - 13	14 - 19	20 - 24	25 - 30	31 - 46

Standard level paper one

Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 3	4 - 7	8 - 13	14 - 19	20 - 24	25 - 30	31 - 46

General comments

The areas of the programme and examination which appeared difficult for the candidates

Understanding the demands of the question continued to be an issue for many candidates. Many candidates did not address the command terms, but simply outlined and evaluated studies. Many candidates also struggled with what is meant by “critical thinking relevant to the demands of the question”. For example, when addressing a question on the role of culture on behaviour, discussing the ethical considerations of a particular study is not highly relevant. Many candidates included a significant amount of irrelevant information, including the aims of the level of analysis, a history of the level of analysis and, in the case of section A questions, unnecessarily evaluating the research. Candidates need to know that these strategies lead to unfocused responses. In addition, there was a tendency to write several examples when only one was asked for. It is important that candidates know that when only one is required by the question, only the first in the response is assessed.

The areas of the programme and examination in which candidates appeared well prepared

Candidates were generally familiar with the syllabus requirements, particular with regard to content. There were some examples of excellent levels of critical thinking and approaches to

addressing the demands of the question. A growing number of candidates are using more modern research and this is a very welcome development.

The strengths and weaknesses of the candidates in the treatment of individual questions

Section A

Biological level of analysis

There were many strong responses to this question. The majority of responses focused on the role of the hippocampus, the frontal lobe or the amygdala. Sometimes, however, the specific part of the brain or its function was vague or incorrect. Some candidates used Sperry's study of lateralization of brain function and thus did not address the demands of the question. If they were able to use Sperry to demonstrate the localization of function for the different hemispheres, this was given credit – but this was not often the case.

Often the study was not described in any detail. It is a concern that many candidates continue to use Phineas Gage without looking at Damasio's work which determined localization of function. The original work by Harlow was not adequate to address the demands of the question. It is strongly encouraged that candidates use more modern research. Brain research from the 19th century is strongly discouraged.

Cognitive level of analysis

There were many weak responses to this question. Many candidates listed all ethical considerations and then carried out a holistic ethical evaluation of a study. Candidates should be encouraged to directly answer the question in the first sentence of the response to the short answer question. Unfocused responses do not earn high marks.

Many candidates did not address the demands of the command term. The question requires candidates to "describe one ethical consideration." Often the ethical consideration was only identified and not described in any detail. For example, for "undue stress or harm", very few candidates described what this meant and often argued that any stress was unethical. There was little knowledge of the role of ethics boards in approving studies, the role of informed consent as a preventative measure or the role of debriefing in guaranteeing that the participants leave the study in "the same condition in which they arrived."

There were many incorrect links of the considerations to the studies – for example, that no consent was obtained in the HM case study; that Loftus did not debrief her participants or that Sacks broke confidentiality rules in the case of Wearing. In addition, several links were superficial in nature – for example, that participants were traumatized by watching the videos in the Loftus and Palmer (1974) experiment.

Sociocultural level of analysis

Many candidates struggled to outline a principle of the sociocultural level of analysis. Many candidates listed several principles and it was difficult to determine which one was the focus of

the response. In addition, many candidates combined two principles – for example, “we have both a personal and social identity so we have a need to belong to a group,” which made both the outline of the principle and the link to the study ineffective. Candidates often did not outline the principle – that is, unpack its meaning in relation to the level of analysis. Studies were often described in good detail, but not always explicitly linked back to the principle.

Section B

Biological level of analysis

There were several very strong responses to this question. Strong responses clearly discussed the interaction between environment and physiology, addressing concerns of how this is measured, individual difference and methodological considerations. Many candidates, however, simply described studies. Often the evaluation of the studies was focused on ethical considerations which were often only of marginal relevance to the demands of the question.

Cognitive level of analysis

Candidates often did not focus on both “how” and “why” the research method is used. In addition, several candidates did not actually discuss the research method, but described and evaluated studies.

It is necessary for candidates to define terms and explain how they are linked to the studies that they are discussing. For example, why is ecological validity a concern in this particular study? What are the concerns about reliability with this particular study? Often critical thinking consisted of formulaic responses which were detached from the study being discussed. In addition, there was often little understanding of the word “control”. Many candidates simply said that experiments are in a controlled environment. There were no examples given of controls used in the studies or why controls were important.

Bartlett was often used as an example of an experiment. This is highly problematic. Arguments were made that it was well controlled, had clear manipulation of an IV, had standardized instructions and participants were randomly allocated to positions. This is not correct. Using old research which did not follow modern standards of research is not a good strategy for addressing this question. It is better for candidates to prepare with research that is more modern and follows clear experimental procedures – for example, Loftus and Palmer. This does not mean that candidates could not achieve top marks with Bartlett’s study. If a candidate described the study as quasi-experimental and how and why this type of experiment would be done, this was a totally acceptable answer.

Sociocultural level of analysis

The vast majority of candidates responded to this question with individualism and collectivism. There were several responses that did not include any description of the dimension, except to say that individualism focuses on the individual and collectivism focuses on the group. This demonstrated limited understanding of the nature of the dimension.

Critical thinking often focused on the studies, rather than on the nature of the dimension. Ethical discussions of Asch, for example, were not relevant to the demands of the question. In addition, candidates often used non-cultural studies (Johnson, Asch) as examples, rather than looking at studies that directly compared cultures.

Many candidates described four to five studies, but the response lacked any real discussion of dimensions. There should be less focus on memorizing a lot of research and more focus on the conceptual understandings demanded by the question.

Recommendations and guidance for the teaching of future candidates

- More time should be devoted to helping candidates master the vocabulary of psychology. Terms such as validity, reliability, credibility, experiment, and correlation were often used incorrectly. Candidates need to know that the misuse of vocabulary influences the quality of their responses.
- More training with the command terms is necessary. Have candidates write responses to short answer questions and extended response questions without any research to see if they can answer the question and include critical thinking. Then go back and have them add examples of research. This should help them to be more focused on the question and less focused on the studies that they would use.
- Many essays were highly redundant. Candidates should plan their responses carefully before beginning to write the response. Make planning a part of in-class assessment.
- Teachers should give feedback to candidates on their work that includes helping them to understand what was unnecessary or irrelevant in their responses – for example, when discussing research on the role of cultural dimensions on behaviour, ethical considerations such as "informed consent" in Berry's replication of the Asch study are only marginally relevant to the question and earn few marks. The command terms should be clearly reviewed by teachers in order for candidates to correctly focus on the requirements of the questions.

Higher level paper two

Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 6	7 - 12	13 - 16	17 - 21	22 - 26	27 - 31	32 - 44

Standard level paper two

Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 3	4 - 6	7 - 10	11 - 12	13 - 14	15 - 16	17 - 22

General comments

The areas of the programme and examination which appeared difficult for the candidates

The overall quality of responses tended to be satisfactory but varied greatly from answers that provided clear and detailed knowledge and understanding relevant to the question to those providing general answers for certain learning outcomes without referring to the specific command term. The majority of answers tended to contain good descriptive knowledge of the required option but failed to address the specific requirements of the question and present a clear argument. Some candidates showed general knowledge of the options, but attempted to make the questions “fit in” with what they had studied. Focus on the question and command term must be obvious in the response submitted, indicating a true understanding of what is being required from the candidate.

Candidates should be continuously reminded that all questions included in paper two require evidence of critical thinking: clear, detailed analysis; relevant discussion of chosen topics, or evaluation of psychological research. Therefore, all attempts to present entirely descriptive knowledge, however detailed, will result in awarding of marks in the lower to middle range.

Answers scoring in the lower ranges had obvious difficulties in structuring a response – poor organizational skills, a tendency toward anecdotal comments or generalized responses lacking in specifics. In addition, many candidates found it difficult to support ideas with relevant psychological research so this was an area that could be improved upon. Some candidates either lost focus during their response, or did not answer what was being asked.

The areas of the programme and examination in which candidates appeared well prepared

Many centres prepared candidates in the area of abnormal psychology. The questions in this option were generally well addressed and it appeared that many candidates were well prepared to respond to these questions in an academic style.

Evaluative skills were demonstrated in the top essays. Methodological and ethical considerations were addressed in skilful ways. Many astute candidates referred to counter claims and studies to reinforce their argument. Reference to psychological research was often provided although precise and focused knowledge of research was not always present. At times, although psychological research was provided it wasn't always effectively used.

The strengths and weaknesses of the candidates in the treatment of individual questions

Abnormal psychology

Question 1

This was not a very popular choice within the option of abnormal psychology. Most candidates chose to write about phobias, obsessive-compulsive disorder (OCD), anorexia or bulimia. Some candidates didn't distinguish different types of eating disorders but rather explained two etiologies of eating disorders in general. Those answering this question tended to do well.

In several cases candidates incorrectly explained etiologies of depression, misidentifying depression as an anxiety disorder rather than affective disorder. This resulted in the loss of marks for knowledge and understanding, and significantly reduced marks for critical thinking and focus on the question being asked.

Question 2

This was the second most popular question within the option. Overall, mediocre to good answers were provided. Most candidates provided good description of biological factors influencing disorders (including information about hormones, neurotransmitters, drugs that help, and parts of the brain responsible for disorders). Although some responses were simply descriptive, others included relevant evidence of critical thinking. Most higher quality responses discussed the extent to which biological factors influence one specific disorder – a very popular choice was depression.

Most higher quality responses tended to focus on more than one biological factor; the effect of neurotransmitters and the role of genes was a popular choice and many candidates provided good responses reflecting relevant, clear and precise knowledge of psychology including reference to relevant theories and studies. A wide range of studies were cited in this response, including Caspi et al. (2003), Neale (2011), Kendler and Prescott (1999), Rampello et al (2000), and Duenwald (2003).

In addition, many candidates provided some additional information about cognitive and/or sociocultural factors in order to respond to the command term "to what extent". This is a good practice in addressing this command term and in providing evidence of critical thinking.

Question 3

Many candidates attempted this question and were able to organize their response well. However, in many responses, the concepts of validity and reliability were not treated as two separate concepts although an understanding of these terms was a key to a successful response as a whole. Candidates were able to provide a reasonable amount of critical thinking skills in response to this question.

Responses to this question offered descriptive (and often inaccurate) accounts of Rosenhan's participant observation (often described as an "experiment"). Good responses included Nicholls's research at Great Ormond Street Hospital and Cooper's study on inter-rater reliability in the UK and USA. Culture, gender, and ethical evaluation tended to be well integrated in the better responses.

High quality responses discussed a smaller number of theories and/or studies in order to demonstrate depth of knowledge. In addressing a larger number of theories and/or studies, candidates tended to be superficial in their description and analysis.

The most glaring problem observed in regard to this question was that instead of addressing the question some candidates chose to discuss concepts of normality and abnormality or instead described classificatory systems in a detailed manner without focusing on validity and reliability of diagnosis.

Developmental psychology

Question 4

This was a very popular question within the option and also within the whole exam. The better responses tended to choose a theory of cognitive development and provided a thorough and thoughtful evaluation. Piaget's theory was by far the most popular and was often well described. Only a few responses discussed Vygotsky's sociocultural theory or biological explanations of cognitive development.

Weaker candidates got caught up in writing pages of description without ever evaluating the theory in any effective manner. While the stages were adequately identified, several important aspects of the theory (such as assimilation, accommodations, egocentrism, etc.) were not always addressed. In some cases, responses were supported with relevant studies but the

relevance of the study was not always made clear. An additional problem with some responses was that the candidates evaluated the studies supporting or disconfirming the theory but not the actual theory.

In some rare cases responses provided evidence of critical thinking by only addressing limitations – in these cases it was difficult for the responses to gain access to higher marks.

Again, weaker candidates tended to not focus on the command term (evaluate) and instead described theories and related studies without regard to the evaluation of said theory/studies.

Question 5

This was a very popular question in this option. Overall the candidates showed very good understanding in their responses and provided good evidence of critical thinking related to theories or studies related to the potential effects of deprivation or trauma in childhood on later development, although in some responses the empirical evidence and/or critical thinking was limited or absent.

Responses reflected a number of different ways in which candidates approached the question. Deprivation and trauma were interpreted in a variety of ways: the case study of Genie was brought up along with descriptions of “wild” children but quite often economic deprivation was also included. Many responses also discussed the longitudinal study of Czech twin boys by Koluchova. Some candidates used Harlow's research on deprivation in rhesus monkeys to very good effect, although not all did a sufficient job of generalizing these or other animal studies (Rosenzweig) to relevant human behaviour. Some strong answers also discussed Hazen and Shaver's research on childhood attachment and its effect on adult romantic relationships. Bowlby's concept of an “internal working model” and Rutter’s research on adopted Romanian children were also presented and some responses offered well-informed discussions of resilience.

Weaker responses did not always address the effect of deprivation and/or trauma on later development, instead focusing upon the immediate effects.

Question 6

This was the least popular choice within the developmental psychology option. Weaker candidates in general were not prepared to discuss the connection between physical change and identity. However, stronger responses made better arguments, citing several studies such as Feron et al (1997), and Simmons and Blyth (1987) which indicate that this is an important developmental consideration.

Health psychology

Question 7

Few candidates attempted this question. The strategies most commonly discussed were taxing certain foods (sugars), dietary changes, and surgery. Candidates, in general, were not able to distinguish between the difference between preventing obesity and treating obesity.

Weaker responses provided superficial knowledge of prevention strategies for obesity with minimal inclusion of knowledge of psychology, instead relying upon mostly common sense knowledge. All in all, this was not a well presented response.

Question 8

Few candidates attempted this question. Those that did presented a wide range of research, including theories/studies relating to stress, health promotion, and/or health problems. However, the description of the research was more prevalent than the evaluation of that research among these responses.

Question 9

This was the most popular question within the health psychology option. In general, most responses provided good understanding of the concepts of problem-focused coping, emotion-focused coping, social support as a coping strategy, mindfulness-based stress reduction (MBSR), etc. Shapiro et al.'s study and Taylor's tend and befriend theory were usually used to support the discussion of MBSR and social support. Most candidates discussed two strategies for coping with stress in great detail.

Several candidates described at length the general adaptation syndrome (GAS) or types of stressors but did not make this part of the response relevant to the question, leading to a very poor answer. Description of strategies such as meditation and yoga were often quite anecdotal. While these techniques are valid, some candidates did not use any psychological research to discuss them which resulted in an answer without much academic merit.

Psychology of human relationships

Question 10

This question was the most popular choice within the option. In many cases the responses to this question received marks in the highest band. Candidates wrote a very well developed and organized response, supported by appropriate theories and studies. However, evaluation for some was quite limited. Latané and Darley's research looking at the role of the number of people available to help (diffusion of responsibility) as well as the informational social influence (pluralistic ignorance) were usually described but not clearly evaluated. Piliavin's cost versus benefit model as well as cognitive dissonance and arousal were also discussed by some candidates. Cultural factors were introduced in the form of evaluative comments. Well prepared candidates evaluated the above theories and research whereas in lower quality essays the theories were described with limited evaluation.

In general, responses that discussed a relatively smaller number of theories/studies on factors influencing bystanderism in greater depth tended to gain more marks.

The biggest problem observed in this response was that some candidates spent too much time describing the Kitty Genovese case in great (and sometimes inaccurate) detail, while ignoring the resultant theories and studies.

Question 11

This question was usually answered well although it wasn't a popular choice. Candidates were well versed in the role of communication in the maintaining of relationships and included reference to evidence from several relevant research studies. The following aspects of communication were frequently addressed: content and amount of communication, self-disclosure, different types of couples (interdependent, independent separate) and cultural differences in communication between couples. Strong answers often addressed Altman and Taylor's research on the importance of self-disclosure and Tannen's research on gender differences in communication.

Some candidates discussed attribution at length, but neglected to explain how attribution affects communication. Several candidates who answered the question on communication seemed to focus on common sense and personal experience, not on psychological knowledge.

Question 12

This question was likewise a popular choice in this option. Most candidates discussed social learning theory and the subculture of violence theory in order to provide sociocultural explanations of the origins of violence. Many of the discussions of social learning theory included weak descriptions of Bandura's "Bobo doll" study. However, higher quality responses sometimes included reference to social identity theory as well.

Other explanations included deindividuation theory (Zimbardo, 1969), Wolfgang's 1967 subculture of violence explanation, and Shanahan and Morgan's 1999 cultivation theory. These were generally better presented and evaluated.

In general, this question was less well answered with candidates often finding themselves tempted to resort to anecdotal explanations of the origins of violence which were not based on psychological theory or evidence. Responses focusing on one or two explanations tended to score higher marks.

Sport psychology

Question 13

This question was the most popular within the option. Many were able to report on research relevant to goal-setting. However, in some cases the role of goal-setting was not made relevant to motivation. Responses commonly discussed goal-setting in relation to performance. Another common problem was that candidates provided descriptive answers that talked only about intrinsic and extrinsic motivation and little else. Responses referring to relevant research studies usually made reference to the following studies:

- Elliot and Dweck (1988) on ego orientation versus task orientation
- Locke and Latham (1981, 2006) on the role of goal-setting in regulating performance and increasing self-efficacy.
- Bandura's (1977) self-efficacy explanation.

Question 14

This question was rather popular within the option and usually responses reflected rather good knowledge of the topic. Responses to this question tended to focus on different aspects of the role of coaches including the following: the role of the coach in regard to the motivation of the athlete; self-efficacy; goal-setting; the role of feedback in improving performance; the role of coaches in team cohesion; and the role of coaches' expectations in the performance of athletes.

Most candidates chose to focus on the positive effect coaches have on the athletes they coach. Some responses included a discussion of the difficulty in isolating variables and the problem of general subjectivity of this type of research. The majority of responses were of mediocre quality.

Question 15

Not enough responses read to provide feedback.

Recommendations and guidance for the teaching of future candidates

- Teaching candidates how to construct an organized response is a big priority and teachers should make sure that all candidates understand how to approach questions, how to effectively deconstruct them and how to structure their response.
- Teachers should encourage the use of terminology relevant to psychology. Many examiners commented that responses were too general and lacked clarity. Providing simple definitions of key terms relevant for the specific question could be a good suggestion for candidates to remind them that all relevant information should be "put on paper" because otherwise it can't be given credit.
- Some candidates did not provide research studies/theories in their responses although this is a general requirement for paper two responses and indicated in the general instructions on the exam paper. Candidates should continuously be reminded to support their arguments with relevant psychological theories/studies.
- It appeared that some candidates had problems in structuring a response that met the requirements of the command term. Teachers should try to focus more on command terms to help candidates apply their knowledge in an appropriate manner. It seemed that the main problem for candidates lay in not being able to interpret the command terms. Therefore, from the very beginning of the course, candidates should be familiarized with the command terms and be exposed to similar kinds of questions as those given in the exam papers, so that candidates are well prepared for the exams.
- Candidates should also be given past paper questions to do and once they are done, the teacher should explain the criteria and markscheme to the candidates. After this, candidates should be asked to assess their work themselves. In this way candidates can take responsibility for their learning.
- Most importantly, teachers should ensure that candidates form a connection between the theory/concept/term and empirical studies and in doing so ensure that they are evaluating the concept, not just the empirical studies, according to the requirements of the question and by referring to the specific command term.

Finally, teachers should not encourage candidates to provide large numbers of studies that are misremembered, and not made relevant. Instead, focus should be made on one or a few detailed studies and one or two updated examples, and then on applying these to answering the question effectively.

Higher level paper three

Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 2	3 - 5	6 - 8	9 - 11	12 - 15	16 - 18	19 - 30

General comments

Overall it appeared that candidates understood the stimulus material quite well and were able to use it to some effect in their responses although there were issues with the command terms in this paper. It was observed that much fewer candidates than in previous years had difficulties with integrating their knowledge of qualitative research methods with the stimulus material and therefore the general impression was that truly generic responses were few.

As usual, some candidates seemed to struggle with things that are not explicitly stated in the stimulus material, for example, arguing that the researcher did not debrief the participants or that reflexivity was not applied in the study and should have been.

There was a slight increase in very good responses indicating that some candidates are very well prepared.

The areas of the programme and examination which appeared difficult for the candidates

Weaker responses to question three demonstrated a tendency to argue based on the assumption that researchers are by nature so biased that they want their own ideas to be expressed in a study. This demonstrates a general lack of understanding that the purpose of qualitative research is to understand the subjective world of participants and that researchers may have a personal interest in a specific topic, as was the case in the study in the stimulus paper, without necessarily losing objectivity if reflexivity is applied.

It seems that some candidates still have a tendency to comment on the study in the stimulus material instead of explicitly addressing the questions related to methodology. Weaker candidates had problems integrating the stimulus material into their response in a meaningful way. Either they relied on heavy citation of the stimulus material in their answer or they treated the stimulus material as a text to analyse rather than using relevant parts of it as support for their arguments on methodological and ethical considerations. Weaker candidates also seem to have problems understanding what is required when the command term ‘explain’ is used as in question one and question three.

In question one, weaker candidates lacked sufficient accurate knowledge of the two interview methods and did not specifically address reasons for choosing semi-structured interviews and

a focus group interview in combination in the study but instead gave a list of strengths and limitations of each method without really explaining how the combination of the two methods could benefit the aim of the study.

Question two seemed to cause problems for some candidates but some very good answers were also seen here. First of all, the command term "discuss" appeared difficult to address effectively in relation to ethical considerations for many candidates. Secondly, it seemed that weaker candidates just listed a number of ethical considerations and failed to discuss why they could be relevant to the study in the stimulus material. Weaker responses claimed that the study was unethical because the researcher had addressed topics considered taboo and consequently did not respect the participants, thereby demonstrating lack of understanding of the requirements of the question.

The areas of the programme and examination in which candidates appeared well prepared

Many candidates used the stimulus material well and could integrate their knowledge of qualitative research methodology with relevant parts of the stimulus material. With reference to question two, a number of relevant ethical considerations and appropriate quotes from the stimulus material were presented in quite a few responses. This indicates that candidates have been well prepared in spotting relevant details in the stimulus material and using them in support of their argument, although the command term ‘discuss’ was often not effectively met.

With reference to reflexivity (question three) many candidates seemed well prepared as they could define reflexivity and argue for use of it in this study with reference to the researcher's own background and current job.

The strengths and weaknesses of the candidates in the treatment of individual questions

Question 1

Question one was often well answered demonstrating accurate knowledge of semi-structured interviews as well as focus group interviews, although there was a tendency to lesser knowledge in relation to the focus group interview. The strongest candidates were able to explain why the researcher had chosen these two methods in her study with reference to method triangulation as a strategy to compensate for potential weaknesses in specific methods. Such responses could, for example, refer to relevant strengths and limitations of the two methods in relation to the socially sensitive topic under investigation, also picking up on benefits of using the participants' mother tongue and the visit before the interview to build rapport as beneficial to the research process.

Weaker candidates had a tendency to give a very superficial description of features of each method and some just listed a brief list of strengths and limitations of each method with no attempt to relate this to the question asked.

Question 2

The command term ‘discuss’ seemed to be a major challenge in this question where many candidates demonstrated accurate and relevant knowledge of ethical considerations. Very few candidates referred to only one ethical consideration even though this was a possibility and an invitation for a more in-depth approach to answering the question. Many candidates referred to a larger number of ethical considerations – sometimes resulting in a rather generic approach instead of focusing on relevant considerations in the actual study.

Candidates generally demonstrated a good knowledge of a number of relevant ethical considerations and listed appropriate quotes from the stimulus material to support their points. Most responses referred to 'informed consent' in line 10 and discussed why informed consent was important in this particular study. Stronger responses also referred to the age of the participants and argued that parental consent in principle was necessary but argued that the importance of the study could justify that this was not obtained.

Candidates also often referred to “protection of psychological harm” by referring to lines 6–9 in the stimulus material.

Weaker responses did not link knowledge of ethical considerations to the study but listed a number of ethical considerations in a quite generic way with no specific reference to the stimulus paper. Some of these did not even refer to "informed consent" in line 10.

Question 3

Most candidates approached this question with explanations of why reflexivity could be important to apply in the study and referred to appropriate points made in the stimulus material such as the fact that the researcher had similar experiences (personal reflexivity) and that she had recruited the participants from her personal network (epistemological reflexivity).

There was no specific reference to the use of reflexivity in the stimulus material and this seemed to have caused some confusion in a number of candidates. Also there seemed to be some confusion about the use of reflexivity as a means to discover possible biases in the research process. Most candidates had a good grasp of reflexivity as well as why it should be used in this study in spite of any specific reference to reflexivity in the stimulus material, for example with reference to the researcher's own background, job role and use of her native language when conducting the interviews, as well as issues of translation.

Few responses demonstrated no clear knowledge of reflexivity in qualitative research or in this study. The weakest responses did not explain what reflexivity is and why it is used.

Recommendations and guidance for the teaching of future candidates

Paper three is based on a short description of a qualitative research study (the stimulus material) accompanied by three questions related to the methodology used in that particular research study. Candidates must answer all three questions paying attention to the command term and using their knowledge of qualitative research as well as information from the stimulus

material to support their analysis. Paper three is based on short answer questions and each question can receive a maximum of 10 marks. Candidates should be trained in addressing each question asked in a straightforward manner and avoid filling in with general knowledge that is not directly relevant to the question asked and will therefore not give any credit.

It appears that candidates this November generally had fewer difficulties using the stimulus material properly but there were still issues. It was often seen that candidates discussed the content of the stimulus material without much reference to relevant knowledge of qualitative research methods. Some candidates wrote about qualitative research methods in a generic way without much reference to the stimulus paper. Consequently, good preparation for the exam involves using past exam papers for training so that candidates will get an opportunity to acquire an understanding of how to apply relevant knowledge and understanding of qualitative research methods to the study mentioned in the stimulus material. It is not necessary to write out long quotes as every fifth line in the stimulus material is numbered so that candidates may refer to the lines and just give a brief outline of the content in these lines in their argument. Teachers could ask students to work with past exam papers and show them how to find relevant parts of the stimulus material that could be used to support their argument with reference to qualitative research methodology.

Likewise, teaching paper three could include exposure to a number of qualitative studies to give candidates more opportunity to understand the philosophy of qualitative research. The optimal strategy is that candidates conduct small research projects on each of the methods in order to get an insight into the reasoning of a qualitative researcher as this would be very useful in relation to developing the thinking skills necessary for paper three.

It is also recommended that teachers provide opportunities to practice the command terms in relation to paper three. Too many candidates still have problems here so understanding what a specific command term requires in paper three should be part of effective teaching.

Finally, it is recommended that candidates are prepared in such a way that they have both (1) a general knowledge of qualitative research methods as outlined in the guide *and* (2) competence in applying this knowledge in relation to the stimulus material as well as (3) competence in using appropriate terms and concepts from qualitative research methods. It is also recommended to train candidates to make balanced evaluations and discussions instead of presenting personal opinions or speculations with limited relevance to the questions asked.