November 2013 subject reports

PSYCHOLOGY

Overall grade boundaries

Higher Level

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 8	9 - 18	19 - 29	30 - 41	42 - 53	54 - 66	67 - 100

Standard Level

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 9	10 - 21	22 - 31	32 - 43	44 - 55	56 - 67	68 - 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 a good standard this session and in line with previous sessions. Almost all internal assessments (IAs) submitted met the criteria for an experimental design, whereby there was manipulation of an independent variable with the effects of a dependent variable recorded. The majority of candidates demonstrated an awareness of ethical issues and included evidence of informed consent, briefing and debriefing instructions in the appendices. There was also an increased awareness of the need to make the background studies and/or theories relevant to the hypotheses.

Most reports were based on cognitive psychology, which lends itself well to experiments at this level of study. Favourite experiments were, as usual, levels of processing, effects of interference and memory recall, reconstructive memory, experiments related to schema

theory, and imagery versus rehearsal. A few candidates performed experiments with several conditions although a simple experiment with ***only two conditions*** is recommended in the guide.

In general, the weaker reports shared the following characteristics:

- Weak and imprecise explanation of background research in the introduction. The hypotheses were therefore not clearly justified. The hypotheses were often not operationalized, that is, made measurable.
- For the descriptive statistics, the use of the descriptive statistics was not explained.
- 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. This was often due to the limited relevant research and/or theories presented in the introduction. Identification of limitations of own procedure was not linked to suggestions for modification.
- Referencing was poor or inconsistent.

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.

Candidate performance against each criterion

Criterion A: introduction

In many IAs, the research presented was generally sufficiently explained and adequate in number. It is, however, still a challenge for candidates to clearly link their research to the hypotheses. Candidates should always *explicitly* state how the theories and/or studies presented in the introduction link 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. For example, rather than 'memory recall will be better in condition A as compared to condition B', the candidate should write (for example) 'the number of words correctly recalled will be significantly higher in the condition with no background noise (silence) as compared to the condition with background noise'.

Criterion B: design

An appropriate experimental design (repeated measures or independent measures) was, for most candidates, correctly identified, but the choice of design was not always properly justified. Candidates should make clear why the particular design was chosen over another.

A number of candidates had problems with operationalization of the IV and the DV, that is, clearly making them measurable.

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 must be stated in this section when discussing ethical considerations.

Criterion C: participants

Candidates continue to neglect to state the target population, that is, the population from which the sample was drawn. Candidates also often confuse the actual sample with the target population.

Most candidates did identify the appropriate sampling technique, but often struggled to explain the use of this method, that is, why the method was chosen.

Overall, most candidates included the relevant characteristics of the participants, although at times irrelevant characteristics were included, for example socio-economic status.

Criterion D: procedure

It is necessary that all materials mentioned (for example, informed consent form, standardized instructions) are referenced in the appendices. Without proper referencing, it would not be possible to properly replicate the experiment.

Criterion E: results – descriptive

Most candidates included a graph and a table, but proper labelling of the graph was still an issue, with incorrect labelling or no labelling at all. At times, it was difficult to ascertain the differences in 'grayscale' on bar charts – it is advisable that candidates print in colour or make the differences in grayscale more apparent.

Often candidates presented the results only in tabular form without describing the results in a narrative form as well.

Only the strongest candidates explained *the use of descriptive statistics*, that is, why the mean or standard deviation was chosen.

Only one measure of central tendency and one measure of dispersion should be presented.

Criterion F: results – inferential

Most 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).

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.

It is also necessary that a conclusion is included for all IAs.

Criterion H: citation of sources

Referencing continues to be an area of weakness. Candidates often did not cite certain research mentioned in the introduction in the referencing section 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 and guidance 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. The manipulation of only two conditions is recommended.

- 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 majority of the reports were satisfactory. The experiments replicated were adequate and often based on cognitive psychology (for example, Stroop effect, the effect of highlighting on memory, weapon focus, Loftus's experiment). A few candidates performed complex experiments with more than one IV and DV – unfortunately their findings and discussion of results lacked precision and thus they did not receive high marks.

In general, the weaker reports shared the following characteristics:

- Weak and imprecise explanation of original study in the introduction (this affected the discussion section as well).
- Descriptive statistics were not appropriately justified or applied. Results were not always focused on the aim of the study.
- Discussions were superficially written and candidates failed to discuss own results in the light of the replicated study.
- Referencing was poor.

Candidate performance against each criterion

Criterion A: introduction

Many candidates presented the key study to be replicated clearly and related this to the aim of the study. However, some candidates merely described certain aspects of prior research in a superficial manner and did not identify the findings and relevance of the study chosen or why they wanted to replicate it. A few candidates provided a general overview of a certain topic in cognitive psychology but failed to clearly identify the study they were replicating.

Criterion B: design

Many reports gained one mark for this criterion. Often candidates failed to justify their choice of design. Also, very often the IV and DV were not operationally defined. In most cases ethical considerations were clearly addressed but at times debriefing

letters were not included in the appendix although candidates stated that participants were debriefed after conducting the experiment.

Criterion C: participants

Again, performance for this criterion was often weak and many reports did not get full marks for this criterion. Candidates often failed to identify relevant characteristics of the sample. Also, the sample was selected most often by using an opportunity sample but this choice was rarely justified in a clear way.

Criterion D: procedure

Most candidates clearly described the procedure of their experiments. Sometimes the description provided lacked relevant information for a clear replication of the study. At times candidates only vaguely mentioned standardized instructions and debriefing letters and unfortunately these were not included in the appendix or they were inadequately written.

Criterion E: results

Some reports lacked clear verbal presentation of results related to the aim of the study. Also, many candidates failed to clearly present their data in tables – often tables only included measures of central tendency while measures of dispersion were either not presented or they were stated verbally with no further elaboration on what these values mean.

In weaker reports graphs tended to be inaccurate or unclear. Candidates should be reminded that for a graph to be understood, they must make sure that those reading it have all the information that they need (the graph's title, legend, axes labelled).

In some cases, candidates claimed that the experiment produced data at nominal level of measurement although this was clearly not the case.

In weaker reports, candidates lacked a clear understanding of descriptive statistics and made an attempt to interpret descriptive statistics by accepting or not accepting the experimental hypothesis.

Criterion F: discussion

Only a few reports gained five or six marks for this criterion. These reports contained:

- a clear interpretation of descriptive statistics related to the aim of the study
- a logical and well presented comparison of findings obtained by the candidate to the findings of the study being replicated

- identification of relevant limitations of the candidate's own work and related suggestions for future modification of the study in the same setting

- an informed conclusion clearly relevant for the aim of the study.

Although most reports succeeded in evaluating their own research not all candidates discussed relevant methodological issues. Also, very often suggestions for future research included irrelevant improvements that did not reflect evidence of critical thinking skills: having more time, obtaining a random sample or having more participants.

Conclusions were usually included but not written with care and precision.

Criterion G: presentation

Presentation in many reports was in accordance with requirements – a lot of reports were impressively presented and scored full marks for this criterion. Unfortunately, there were still reports in which the maximum number of words permitted was exceeded or in which no reference section was provided.

In some cases, although standard citation methods were used there was a lack of in-text citations.

Another common mistake was failure to label appendices and reference them in the body of the report.

Also, raw data were sometimes incorrectly included in the results section.

Overall in the majority of the reports the presentation was reasonable. However, care needs to be taken by candidates to ensure that the information presented in the method section is entered under the appropriate subheadings.

In some weaker reports abstracts were poorly written and did not contain all relevant information.

Raw data were not always included in the appendix or presented with calculations of descriptive statistics.

Recommendations and guidance for the teaching of future candidates

- The precise nature of what is necessary for a study to be considered an experiment should be examined by teachers and its correct meaning clearly explained to all candidates. It is recommended that candidates replicate experiments in which **one** independent variable is manipulated and the effect of this is measured on **one** dependent variable. When multiple variables are used (manipulated or measured) this almost inevitably leads to confusion and the use of longer, less clear

explanations.

- Candidates should provide a thorough explanation of the original study – this explanation needs to include not only the aim and the findings but also relevant procedural aspects of the original study. This information should be selected and presented wisely so that it can be used later in the discussion section when a comparison of the two studies is provided.

- A clear justification of the sampling method should be included in the participants section. Candidates should also describe relevant characteristics of participants – the characteristics presented should be important for their study and also include precise information such as age, sex, nationality or cultural background, knowledge of English.

- Only one measure of central tendency and one measure of dispersion should be included for criterion E. Candidates need to practise identification of optimal measures depending on the type of data gathered. Teachers should encourage candidates to check their calculations, clearly write them and include them in the appendices. For full marks for criterion G, calculations should be included in the appendices.

- The explanation of the replicated study from the introduction must be referred to in the discussion section. New studies or theories should not be introduced here. Candidates should be trained in writing a discussion section (perhaps by reading a couple of research articles to become familiar with the idea and style) and they should consult the checklist to be sure that all the IB requirements are met, for example in terms of discussing statistics from the results sections. Understanding of relevant limitations of own research and suggestions for modification should be tied together and it is not enough to say that a particular study should use random sampling and more participants to be valid. The limitations should be stated clearly and they should be linked to the candidate's own experiment.

Higher and standard level paper one

Component grade boundaries

Higher level

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 4	5 - 9	10 - 12	13 - 18	19 - 24	25 - 30	31 - 46

Standard level **Grade:**	1	2	3	4	5	6	7
Mark range:	0 - 4	5 - 9	10 - 12	13 - 18	19 - 24	25 - 30	31 - 46

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

Understanding the demands of the question continues to be an issue for many candidates. Many candidates employed a strategy of listing responses or simply describing a story, rather than addressing the demands of the command term. 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, unnecessarily evaluating the research. Candidates need to know that these strategies lead to unfocused responses. In addition, there is a tendency to write several examples when only one is 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 essays often had long, unfocused introductions. Often the focus of the essay was on a description of many studies. Candidates may do well by using only two to three studies, but demonstrating critical thinking and linking the studies back to the demands of the question.

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.

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

Section A

Biological level of analysis

Many candidates struggled to choose an appropriate study. Many candidates chose the case study of David Reimer, HM or Clive Wearing, but were unable to link it to biological research or to describe ethical considerations. Many candidates also chose Schachter and Singer which is a study of cognitive appraisal. This study earned no credit. It is important that candidates outline the aim, procedure and findings of research and that they make clear links to the level of analysis.

Many candidates simply listed the ethical considerations and did not describe them. Strong responses described the ethical consideration in detail and linked it to the study. There were many superficial responses, for example, arguing that animals cannot give consent, or incorrect descriptions of studies, for example, Caspi revealing to a participant that they could get depression.

Cognitive level of analysis

There were some very good responses to this question; however, many candidates simply described a study of the reliability of memory but did not explain it. For example, in describing Loftus & Palmer's (1974) study, it was necessary to link it back to schema theory as an explanation of why the memory was unreliable. Often the demands of the command term were not addressed.

Sociocultural level of analysis

There were several excellent responses to this question. Candidates demonstrated sound understanding of compliance techniques and were able to give examples of appropriate research. Once again, however, several responses did not address the command term "explain." Often a principle was stated, such as "commitment" or "reciprocity", but no actual explanation was provided.

Section B

Biological level of analysis

Overall, candidates demonstrated a clear understanding of the differences between structural and functional imaging techniques. There were good descriptions of the technology and examples of research were, for the most part, relevant and well described.

However, critical thinking related to this question was often superficial. Many of the responses focused on ethics, whether participants had metal implants, or the cost. Very few candidates demonstrated the actual concerns about using imaging techniques in biological research – that is, the distortion of data, the correlational nature of much of the research, ecological validity or the role of confounding variables. When evaluative points were made, they were often stated and not developed or linked to the research being described.

Cognitive level of analysis

Many candidates chose this question and there were quite a few very strong responses. Strong responses explicitly identified the social and/or cultural factors and discussed their effect on a cognitive process. These responses addressed the role of cultural norms, education, environment (resources) or cultural dimensions. Several candidates, however, wrote essays which evaluated schema theory, which was not relevant to the demands of the question. There were also several responses that described several studies. Fewer studies that are discussed using critical thinking would result in a higher quality of response.

Sociocultural level of analysis

Many candidates presented an accurate discussion of how and why a relevant research method, such as experiment or observation, is used at the sociocultural level of analysis. However, some candidates presented only different ways in which an observation is done, such as covert observation, without addressing the key features of the method per se. Moreover, many candidates evaluated the selected method rather than discussing how and why the method is used

There were several candidates who described Asch, Bandura, Zimbardo or Milgram's experiments, but were not able to explain how the study as described was actually an experiment. There were several incorrect responses that referred to these studies as observations.

Recommendations and guidance for the teaching of future candidates

- Teachers need to focus on developing essay-writing skills. It is important that candidates realize that there is no single strategy that will allow them to answer all of the questions. 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 how social and cultural factors affect memory, ethical considerations such as "informed consent" 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.

- Spend time helping candidates to "prioritize" information. Candidates often include a lot of details which although interesting, are not highly relevant to the demands of the question. For example, in the Bouchard study of intelligence, information about the personal lives of specific participants is not relevant to the question. In addition, have candidates brainstorm the studies that they could use and then prioritize the list of studies, justifying why the "number one study" deserves that position on the list. Choosing better examples to respond to the questions, as well as taking time out to plan the essay responses before writing, will help the candidates to write stronger arguments.

Higher and standard level paper two

Component grade boundaries

Higher level							
Grade:	1	2	3	4	5	6	7
Mark range:	0 - 4	5 - 9	10 - 15	16 - 20	21 - 25	26 - 30	31 - 44

Standard level **Grade:**	1	2	3	4	5	6	7
Mark range:	0 - 2	3 - 4	5 - 7	8 - 10	11 - 12	13 - 15	16 - 22

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

Most of the candidates understood the general requirements of the questions set for paper two. However, some candidates had difficulty in addressing the command terms which made the task of handling the questions with appropriate knowledge difficult. This concern was often related to responses to questions which required the candidates to 'contrast' and address 'to what extent' are certain factors related. Overall evidence of critical thinking for most of the responses tended to be rather limited or weak and explanation of concepts, theories and links was often basic. In addition the majority of candidates found it challenging to direct their critical thinking towards the main focus in the question, instead of routinely evaluating empirical studies related to the topic. Forming a connection between the empirical study and the specific topic of the question was often performed in a superficial manner.

Responses awarded marks in the lower ranges had obvious difficulties in structuring a response – poor organizational skills, a tendency toward anecdotal presentation of information or generalized responses lacking in specifics. In addition, some candidates found it difficult or forgot to support ideas with relevant psychological research.

Some additional comments related to specific questions:

- Some candidates who answered question one discussing concepts of abnormality failed to include psychological research in their response.
- Some candidates had difficulties focusing and organizing their response to question seven and often provided long introductions in which they described stress and explained how stress has a negative impact on human health and behaviour. Although this information could be part of a relevant discussion it was usually too long and wasn't well linked to strategies for coping with stress.

- As in past examinations, the command term “evaluate” (in questions 5, 7 and 12) often led to descriptive accounts with minimal efforts to provide a pertinent evaluation.

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

Overall most candidates were able to maintain the structure of the essay and provided some introduction, main body and conclusion although the focus of the response sometimes kept shifting. Most responses indicated very good knowledge of the main concepts relevant for the option. For the most popular options (namely: abnormal psychology, developmental psychology and psychology of human relationships), the responses were very well organized, with some argumentation present. The majority of responses clearly provided empirical evidence.

In centres that were well prepared for the examination the questions were clearly addressed, often beginning with defining key terminology that helped in structuring the response. In these cases research studies and theories were used efficiently in order to support the key points of the response.

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

Abnormal psychology

Question 1

Many candidates attempted this question and in doing so they were able to organize their response well. Most candidates were able to provide a reasonable amount of critical thinking skills in response to this question, including candidates who were not able to score marks in the highest band. Cultural and ethical evaluation tended to be well integrated in the better responses.

The best answers referred closely to research and discussed several concepts of abnormality, remembering to also include the reasons why it is important to consider them.

However, often responses to this question offered descriptive (and often inaccurate) accounts of Rosenhan’s participant observation (often described as an "experiment").

The biggest problem present in some weak responses was that instead of addressing the question some candidates chose to describe classificatory systems in a detailed manner without focusing on concepts of abnormality. Also, some candidates provided long descriptions of several concepts of abnormality (most popular were: statistical infrequency, deviation from social norm, maladaptiveness) but failed to provide clear evidence of critical thinking and forgot to include some empirical or theoretical support thus writing the answers in a superficial manner.

Overall good general knowledge was provided but not many excellent answers.

Question 2

Ethical considerations in diagnosis were usually well addressed. Issues discussed were: misdiagnosis, stigmatization, labelling, self-fulfilling prophecy, diagnosis leading to treatment which may have unwanted side effects.

One weakness noticed in many responses was that even though some candidates put much effort into the description of Rosenhan's research, there were weak links provided to ethics of labelling in diagnosis. Also, Rosenhan's study was often described in an inaccurate manner. Another study candidates had problems with was the study conducted by Langer and Abelson (1974) – again inaccurate descriptions were provided.

Cultural considerations were sometimes included as part of the response. Candidates referred to the following issues: cultural bias or misdiagnosis due to lack of understanding of patient's cultural background. Well prepared candidates managed to link this information to the requirements of the question but some candidates failed to provide this link and therefore the information provided seemed irrelevant for the question asked.

The biggest problem present in weak responses was that instead of addressing the question some candidates chose to discuss ethical considerations in general or ethical considerations related to research studies.

This question was not very popular but in the majority of cases candidates choosing this question had relevant knowledge of the option. Some excellent answers were provided that included a balanced review of a few ethical considerations in diagnosis.

Question 3

This was the most popular question within the option. Overall, mediocre to good answers were provided. The majority of candidates who attempted this question were able to provide some descriptive knowledge while higher quality responses thoroughly contrasted two approaches to treatment.

Good responses usually identified a specific disorder (most candidates chose to write about depression but a few decided to write about anorexia, bulimia or phobias) and then described and discussed two approaches to treatment of the chosen disorder supporting their responses with reference to relevant empirical studies. The best responses usually dealt with depression and chose drug treatment and cognitive-behavioural therapy.

Lower quality responses focused too much on describing the chosen dysfunctional behaviour and offered general comments about one or both approaches to treatment. Often more emphasis was put on drug treatments. In these cases cognitive-behavioural therapy was described superficially and in some cases it was impossible

to recognize if the therapy was described as an individual or group approach. Also in lower quality responses candidates provided some evaluation of the treatments but failed to contrast them and failed to provide research support.

Some candidates did not address the question as it was stated but gave a general and vague description of one anxiety, affective or eating disorder and suggested and outlined two possible treatments, ending the response by suggesting that an eclectic approach would be the best solution.

Developmental psychology

Question 4

This was a rather popular question. Very often this question was answered in a broad and general manner by providing many factors that may affect cognitive development – most candidates chose to write about the relevance of parenting, diet and culture for cognitive development. The majority of responses provided overly descriptive accounts of relevant empirical studies including some evaluation of these studies but failing to provide a clear discussion of the factors.

Weak responses to this question offered descriptive and general accounts about the relevance of diet without linking this specifically to cognitive development.

Some candidates misinterpreted the question and instead of discussing how social and/or environmental variables may affect cognitive development they discussed how lack of attachment influences later social and emotional development. Other candidates gave an overview of Piaget's or Vygotsky's theory with a lot of detail but this information was largely irrelevant for answering the question being asked.

Question 5

Very few candidates took up the task of writing a response to this question. Most responses focused on Erikson's theory of adolescence or Marcia's (1966) studies on the different types of identity status. Understanding and knowledge of these was rather good but evaluation tended to lack depth.

In a few weak responses there was an overemphasis on Erikson's theory of overall development including all stages of development from birth till death.

Question 6

This was a very popular question in this option. Overall, candidates showed very good understanding in their responses and provided some evidence of critical thinking related to potential effects of deprivation or trauma in childhood on later development.
Responses reflected a number of different ways in which candidates approached the question since deprivation and trauma were interpreted in a variety of ways: the case study of Genie was often provided with other descriptions of case studies of deprived children but quite often economic deprivation and poverty were also included.

Bowlby's theory of attachment and Rutter's research were also often included as well as informed discussions of resilience.

Most responses identified and described psychological research related to deprivation or trauma but in some responses these presentations of case studies or experiments were too long and inaccurate. Some candidates referred to animal studies as part of their response, but they failed to link the findings to human behaviour.

Health psychology

Question 7

This was the most popular question for this option and in the majority of cases it was attempted by candidates who had good knowledge and understanding of this option. Most candidates were able to outline or describe two strategies for coping with stress. Popular choices were: stress inoculation training (Meichenbaum, 1985), social support groups/networks, mindfulness-based stress reduction strategies, alcohol abuse, smoking and over-eating.

Evaluation was often presented in a rather simplistic way by stating that one strategy is better or that one strategy tends to have a long-lasting effect in comparison to the other strategy.

Overall, good knowledge was present but there was limited evidence of critical thinking. The question allowed for clear structure in the response that was easily supported by relevant research.

In some responses, too much time and effort was invested in answering what stress is and what effect it has on people, which resulted in accounts that were too short and simplistic evaluation of strategies.

Question 8

This question was a rather popular choice but most responses provided a general and vague account of some sociocultural factors – most popular choices were socio-economic status, poverty, sedentary lifestyle and culture. Most responses provided superficial knowledge and understanding lacking empirical support and were not able to clearly address the "to what extent" command term.

The best responses addressed the question by offering a pointed analysis and evaluation of a limited number of sociocultural factors and the extent to which they influence one specific health-related behaviour (popular choices were obesity and substance abuse).

Question 9

This question was not a popular choice within the option. The most popular choices of models/theories were: the health belief model (HBM), stages of change model, theory

of reasoned action and social learning theory. Most responses provided a good description of one or both models and provided some empirical support for these models but often failed to discuss them in a thorough manner.

Some candidates spent too much time and energy focusing on one model/theory related to health promotion and failed to address the second one in sufficient detail and depth.

Psychology of human relationships

Question 10

This question was very popular within the option. Most of the candidates were able to provide an accurate and elaborated explanation, forming a connection between sociocultural factors and its effect on human relationships. A number of different sociocultural factors were addressed. Most popular choices included: proximity, media influence on violence or prosocial behaviour, culture of honour, the role that culture plays in the perception of attractiveness and the effect of culture on formation and maintenance of relationships. Usually the empirical evidence provided was sufficient to answer the response.

A few responses showed only superficial knowledge of the link between sociocultural factors and human relationships.

Question 11

There was a mixed range of responses to this question, including some excellent responses and some very rudimentary responses.

Most candidates chose to write about the effects of exposure to bullying and terrorism. The effects most often discussed were: physiological responses to stress (for example, fight or flight response), lower performance in school and effects on mental health (for example, anxiety, low self-esteem, PTSD, depression, suicide).

Excellent responses provided a balanced review including empirical evidence in support of the notion that violence produces short-term and long-term effects but also giving a counter argument by referring to presence of resilience in some cases.

Question 12

This question was a rather popular choice. Some candidates wrote a very well developed and organized response providing good knowledge and understanding of relevant studies and/or theories. Popular choices were: Bradbury and Fincham (1992), Tannen (1990) and social penetration theory. However, evaluation tended to be limited and was usually focused on methodological considerations of studies or strengths and limitations of chosen theories. Overall, mediocre to good responses were provided.

Some candidates discussed the role of communication in maintaining relationships in a vague and general manner but made no reference to psychological research. This approach resulted in very low marks.

Sport psychology

Question 13

Only a few candidates attempted this question and the answers did not cite any supporting research and seemed based on personal experience rather than the option having been taught or researched.

Question 14

In the answers provided there was a lack of supporting research. Only a few answers showed real critical thinking skills.

Question 15

The answers provided were very simplistic and were based on experience and common sense. There was little supporting evidence provided making answers appear quite superficial.

Recommendations and guidance for the teaching of future candidates

- It seems that many candidates still have problems interpreting and addressing the command terms in a focused manner. From the very beginning of the course, candidates should be familiarized with the command terms and be exposed to them regularly.
- Candidates must be taught how to break down the requirements of a question and how to effectively structure a focused response and support their arguments with psychological research.
- Specific psychological terminology and concepts must be taught.
- More emphasis should be put on evaluating studies and theories in a meaningful and thorough manner. Instead of providing a large number of studies that are sometimes inaccurately described a better strategy for candidates would be to focus on one or two studies or theories and explain their relevance for the question stated in a thorough manner.
- In centres where teachers are new to the IB programme it is imperative that teachers are aware of the rigour and demands of the IB psychology programme. Generalized, anecdotal discussions are not acceptable responses to examination questions.

Higher level paper three

Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 2	3 - 4	5 - 7	8 - 10	11 - 13	14 - 16	17 - 30

General comments

Most candidates seemed to have a very good understanding of the stimulus material. Generally candidates used the stimulus paper in a relevant way but as usual some candidates had problems integrating the stimulus material with knowledge and understanding of qualitative research methodology. The purpose of paper three is to demonstrate knowledge of qualitative research methods as well as how to apply it to the stimulus material as outlined in the guide and overall, many candidates demonstrated that they knew this well. Only a few candidates did not refer to the stimulus paper at all in their responses. It was evident that even weaker candidates knew that they had to combine knowledge of qualitative research with the stimulus material. However, weaker candidates simply did not have sufficient knowledge of qualitative research methods to address the questions appropriately. The weak responses used various strategies for answering the questions – for example, citing heavily from the stimulus material or writing what they knew of qualitative research. Some of the weaker candidates also addressed obesity as the problem in the stimulus material rather than focusing on the problems doctors faced when treating obese patients.

Another thing observed in November was a problem with addressing the command term – especially in questions one and three.

As in previous years, quite a few candidates used the term 'experiment' interchangeably with 'research study' but it did not affect their reasoning on qualitative research methods.

There was some spread in the marks awarded. Candidates scored all along the mark range with some in the low range, most in the middle and fewer in the higher range. This is an indication that many candidates are well prepared to answer paper three questions.

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

The most difficult question for candidates appeared to be question three where they had to discuss the extent to which findings from the study could be generalized. Weaker candidates mostly referred to parts of the stimulus paper but did not really connect this to knowledge of generalization from qualitative research. Or candidates had no knowledge of generalization from qualitative research and only referred to statistical generalization.

Question one presented a problem to weaker candidates who demonstrated no or very limited knowledge of considerations before conducting the interviews, or they wrote at length about things of marginal relevance to the question.

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

Many candidates demonstrated sound knowledge and understanding of inductive content analysis. In many cases there was also appropriate reference to themes in the stimulus material.

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

Question 1

There were many good answers to this question. Overall, this question was often well answered and strong responses demonstrated that candidates could apply their knowledge on qualitative research methodology to the stimulus material. Stronger responses typically referred to sampling, choice of interview method, choice of interviewer and training, data recording techniques, choice of method of transcription and ethics, and they explained why the chosen considerations were relevant in the context of the study.

The weaker responses had difficulties referring to the stimulus material, were quite generic and did not meet the demands of the command term. Some of the weaker candidates only referred to ethical considerations and often in a generic way. Other weaker candidates had problems being explicit about considerations *before* the interview. Others only evaluated the use of the semi-structured interview and therefore did not really answer the question.

Question 2

There were some really good responses that scored in the highest markband – but also a number of weaker ones. Description of the process of inductive content analysis (ICA) was more or less detailed with reference to the themes mentioned in the stimulus material. Most candidates were able to identify the common themes from the stimulus material and could use that but sometimes this was the only relevant point made in the response. Stronger responses gave a detailed description of the steps in ICA (for example, transcription of the interview, read and re-read the transcripts to identify categories, coding, identification of lower-order/higher-order themes, summary table and so on) with relevant examples from the stimulus material. Such responses often also included description of various strategies for controlling interpretation of the data.

The weaker responses demonstrated limited knowledge of inductive content analysis. In these responses description of the procedure of ICA did often not go into any depth

and many responses went off track and actually evaluated the use of ICA instead. Some candidates gave detailed descriptions of various ways to transcribe interviews but this was only marginally relevant as the focus of the question was ICA.

Question 3

The question seemed difficult to some candidates. Weaker responses typically only referred to the type of sample as a criterion for generalization, for example with reference to culture, gender, age and location of the doctors as determining the degree to which generalization could be made to a larger population or worldwide. This showed limited knowledge and understanding of generalization in qualitative research. However, some candidates seemed very well prepared and were able to either discuss generalization of findings from the study in the stimulus material, for example referring to collection of rich data and comparing to findings of other similar studies. The stronger candidates were also familiar with the terminology related to generalization in qualitative research, that is, inferential, representational or theoretical generalization and responses in the higher scoring range could also discuss this in relation to the stimulus material.

Recommendations and guidance for the teaching of future candidates

- Each exam paper is based on a brief description of a qualitative research study (the stimulus material) combined with three questions. All questions must be answered. The stimulus material is intended to serve as a starting point for analysis of how qualitative research methods could be applied to a specific study.

- It is recommended to base teaching of this part of the programme on practice of "what it is like to be a qualitative researcher", for example, having candidates undertake activities that enable them to reflect on various aspects of qualitative research methods. Teaching paper three should include exposure to a number of qualitative studies to give candidates the 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. It is equally important that candidates have trained with previous exam papers so that they become familiar with the requirements of this paper.

- Using previous exam papers gives candidates an opportunity to understand how to apply relevant knowledge and understanding of qualitative research methods in the context of the stimulus material. Every fifth line in the stimulus material is numbered so that candidates may refer to the lines without having to use extensive quotations. This could be used more effectively in the responses and teaching this paper should involve showing candidates how to find relevant parts of the stimulus material that could support explanation or discussion of qualitative research methodology.

- Overall, candidates should be 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. It is also recommended to train candidates to make balanced evaluations and discussions instead of claims and speculations with limited relevance to the questions asked.

- Finally, it is advised that teachers instruct candidates in what it means to address the command term in relation to paper three, for example what "explain" or "discuss" means. Too many candidates still have problems with this so understanding what a specific command term requires should be part of effective teaching.