May 2015 subject reports

Psychology TZ2

(IB Africa, Europe and Middle East and IB Asia-Pacific)

To protect the integrity of the examinations, increasing use is being made of time zone variants of examination papers. By using variants of the same examination paper candidates in one part of the world will not always be taking the same examination paper as candidates in other parts of the world. A rigorous process is applied to ensure that the papers are comparable in terms of difficulty and syllabus coverage, and measures are taken to guarantee that the same grading standards are applied to candidates' scripts for the different versions of the examination papers. For the May 2015 examination session the IB has produced time zone variants of psychology paper one.

Overall grade boundaries

Higher level

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 8	9 - 18	19 - 30	31 - 42	43 - 54	55 - 66	67 - 100

Standard level

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 8	9 -20	21 - 31	32 - 43	44 - 54	55 - 66	67 - 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, with a clear understanding of the requirements of the internal assessment (IA). As in last year's May session, there was a notable reduction in IA work that did not meet the criteria for an experimental study. The majority of candidates were aware of ethical issues and included a copy of the informed consent, briefing and debriefing instructions in the appendices.

Most reports were based on studies from cognitive psychology and this seems to provide good results at this level of study. Favourite experiments were, as usual, reconstructive memory and experiments related to schema theory, studies related to the duration of the short-term memory, stroop effect, and imagery versus rehearsal.

Some issues with IAs still persist such as:

- Three or four variables being manipulated although a simple experiment with only two conditions is recommended in the psychology guide.
- For the descriptive statistics, the use of descriptive statistics was not explained.
- For the inferential statistics, tests were identified but not justified. Also, many candidates failed to include the raw data or calculations of the inferential test chosen.
- Discussions were superficial with limited discussion of the IA results in the light of background research and/or no reference to statistics. Identification of limitations of own procedure was not linked to suggestions for modification.

It should be noted that it is not required to make an exact replication of an experiment. A partial replication is adequate but the candidate's experiment should be closely linked to an actual experiment.

Candidate performance against each criterion

Criterion A: introduction

In some IAs, the research presented was not *explicitly* linked to the hypotheses. Candidates should always clearly describe the research (theories and/or studies) and state how they link to the hypotheses. Contradictory research should not be presented.

It is important that the background research presented logically leads towards the research hypotheses and the background research is explained and analysed in sufficient depth. This allows for the formulation of a clear research hypothesis that in turn will stimulate discussion of results in the light of the background research in the discussion section.

Candidates seemed to have difficulty clearly writing the hypotheses. The variables should be operationalized and the wording should be clear as to what the expected outcome will be.

Criterion B: design

Most candidates stated an appropriate design (repeated measures or independent design) but the choice of the design was not always properly justified (that is, why that 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).

Criterion C: participants

The target population, that is, the population from which the sample was drawn, was not always appropriately identified. Often candidates confused the actual sample with the target population.

Overall, most candidates included the relevant characteristics of the participants, such as age, gender, colour-blindness (in the case of the Stroop Effect) and/or English level proficiency.

The sampling technique was largely correctly identified, but the use of the technique was often not explained.

Criterion D: procedure

Candidates should make sure to make reference to all ethical guidelines that were followed. It is also necessary that all materials 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, with sufficient labelling. Most candidates also included a description of the results. Only one measure of central tendency and one measure of dispersion is required.

As in previous sessions, only the strongest IAs explained the use of descriptive statistics, that is, why the particular measure of central tendency and dispersion was were chosen.

Criterion F: results – inferential

Most candidates did choose an appropriate test and did justify 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. There seemed to be an increase this session in the number of candidates who did include the raw data or the inferential test calculations. It is important that raw data and all calculations of the inferential test are included in the appendices. If the calculation is performed online, a screen shot of the calculation could be included in the appendices as documentation.

A number of candidates did not make a statement of statistical significance and/or the null hypothesis was not accepted or rejected, which is required for full marks.

Criterion G: discussion

As with previous sessions, this section in the report seemed to present the most difficulty for the candidates, as it often lacked development and analysis. The results of the IA were often only referred to but a discussion is required. Candidates should always refer back to *all* research presented in the introduction and discuss these in reference to their own findings.

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 a general experimental nature. There is no need to include the strengths of the design and procedure.

It is also necessary that a conclusion is included.

Criterion H: citation of sources

Candidates often did not include references for research mentioned in the introduction. Additionally, candidates did not use a standard citation method, such as APA, or referencing was not complete.

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.

Recommendations and guidance for the teaching of future candidates

- It is recommended that teachers help candidates find appropriate background research, that is, a theoretical framework and appropriate studies. Finding relatively simple experiments to replicate is recommended.
- The background research in the introduction should be analysed in sufficient depth so that the aim of the candidate's own research is clearly justified and the experimental hypothesis should be clearly linked to background research.
- For the sample, the number of participants in the experiment does not need to exceed 20 (independent design) or 10 (repeated measures design).
- It would be helpful if candidates were given past experiments to read in order to familiarize themselves with the aspects of experimental research. Some candidates would benefit from doing a "pilot IA" in order to familiarize themselves with the format and procedure of an experimental design.
- Candidates should be taught how to properly reference research, as often the citation of sources was incomplete or inconsistently presented.

It is generally recommended that candidates are familiar with scientific standards, and the reading of proper background research should be encouraged. It is recommended that candidates be trained in critical use of internet resources. Many candidates only used internet sources of a non-specialist nature as background literature.

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

A wide variety of work was submitted. Almost all work was focused on appropriate topics and used appropriate designs. Cognitive psychology remains the most popular area of research for IA reports, while a few were also from social psychology.

Most centres chose a manageable research topic that required an appropriate manipulation of an IV and a DV. A few centres allowed candidates to conduct experiments on conformity, which is not allowed.

There were a surprisingly high number of reports of studies where candidates failed to manipulate an independent variable. For example there were studies that used left- and right-handedness, or gender, as the independent variable. Any pre-existing condition cannot be used as a basis for difference in the two groups. All independent samples designs must be able to use random allocation to the groups or provide evidence of two clearly different experimental procedures in order to qualify as an experiment for the purposes of the IB psychology internal assessment.

There were quite a few reports that qualified as poor experimental procedures. These were typically serial position or schema processing where participants were tested only once, the design was said to be repeated measures, and two different variables were included in the test. However, this approach does not allow for random allocation to conditions and/or there is no presence of two different conditions presented to the same or different participants. Many serial position effect studies were in fact surveys as the independent variable was not manipulated.

Some studies involved designs that were more complex than required.

The range of work differed mainly in the introduction and discussion sections. In the introduction section some papers lacked an in-depth analysis of previous research but rather provided a summary of several studies. Other papers clearly explained the original study at an in-depth level. In the discussion section some papers did not connect or analyse the findings in a meaningful way, but rather the discussion section was written on a very superficial level.

Candidate performance against each criterion

Criterion A: introduction

The background study that the candidates identified was usually well related to their work. The aim of the study was not always stated clearly in terms of the IV and DV. Instead some candidates stated that their aim was to replicate a previous study. Occasionally a hypothesis was stated rather than the aim.

The majority of candidates were well-acquainted with the findings of the replicated studies, but many candidates forgot to state all relevant aspects of the original study (aim, type of design used, target population and sampling method, procedure, results obtained and conclusions reached).

Occasionally moderators noticed that candidates were attempting to write a review of literature but failed to identify and explain in detail one study that had actually been replicated.

Criterion B: design

The design was usually well handled by the candidates; however, a number of candidates did not clearly justify the design, but stated what is meant by an independent samples design. Justification of the design should include a rational explanation on the strategic advantage of one of the strengths or characteristics of the design chosen.

Designs were adequately chosen and competently justified in higher quality reports, with variables being precisely defined. Ethical principles were followed in a disciplined way. Candidates should understand and more clearly apply the importance and characteristics of informed consent to participate as opposed to simple consent to participate.

Criterion C: participants

A number of candidates failed to identify and/or justify their sampling procedure. Additionally, a number of reports did not specify the target population or its relevant characteristics.

Criterion D: procedure

Procedures were sufficiently described to allow replications. All materials should also be presented in the appendices to allow for full replication. If particular materials or stimuli were selected or developed, this should be noted and justified.

Criterion E: Results

The responses in this section varied substantially according to the marking criteria.

Most candidates seemed to have a sound understanding of the descriptive results required in this section. However, many candidates are not taking the advantage to fully discuss the meaning of the results of the descriptive statistical tests – both central tendency and dispersion. For example, describing the difference in the values of two standard deviations of the experimental and control group will allow for some great topics in the discussion section. A

rather large number of reports included no measure of dispersion or included a measure of dispersion that didn't fit with the measure of central value.

A few reports employed inferential tests that were not necessary for standard level IA reports. Complex designs contributed to this problem, as candidates did not know how to deal with the huge amount of data generated by the design.

Criterion F: discussion

The standard of this section varied substantially amongst candidates. A number of candidates were well aware of the large mark allocation to this section and dealt appropriately with their findings in light of the original study. They were also able to identify limitations of their research and develop logical suggestions for future studies. Unfortunately, a number of candidates struggled with this section, often failing to relate meaningfully to previous research, or just briefly mentioning the findings of the original study. All candidates seemed to identify limitations of their research but struggled to identify limitations which would be relevant for the design or procedural aspects of the experiment.

Discussions need to document candidates' competence in both interpreting the findings and addressing methodological issues. In a few reports the conclusion was either not presented or not relevant to the aim of the study.

Criterion G: presentation

References need to be written in a consistent and standard manner; this includes alphabetical order. It is important to state the word count on the front page and stay within the word limit, as clear and precise writing skills are important. In some weaker reports the abstracts were written poorly and were not complete.

Recommendations for the teaching of future candidates

In general, candidates presented well-researched and well-thought-out experiments.

It is recommended that candidates choose a simple topic with a clear manipulation of the independent variable and that they choose a clear dependent variable that can be easily measured. The operationalization of variables seemed to be one of the biggest problems with candidates who seemed unaware of this relevant decision.

Focus on different sections tended to vary considerably from centre to centre. Overall the discussion section is commonly where candidates should improve and invest more time and knowledge. A greater understanding of the psychological terms, that is, different forms of validity and reliability would help candidates in their evaluation of research.

The design of experiments that will both satisfy the IA requirements and generate data that can be analysed appropriately by candidates needs a more thorough approach. The choices and description of descriptive statistics and verbal account of tables and graphs need more guidance. Candidates should be encouraged to reflect on weaknesses of their studies, provide

a meaningful link between their findings and the findings of the original study they base their work on.

Teachers should include some simple exercises during the course of the first year that will help candidates with writing the introduction and discussion section.

The methodological component of the course may need to be introduced earlier, so that candidates understand the meanings of the terminology within the marking criteria: target population; sampling techniques; justification for using a particular sampling technique; descriptive statistics.

Practice in using descriptive statistics and choosing the most appropriate measure for the design and type of data may improve the quality of the report. It would be useful to focus on the strengths and weaknesses of different sampling methods and designs and encourage candidates to justify their decision of a particular design or sampling method.

Higher and standard level paper one

Higher Level Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 3	4 - 7	8 - 12	13 - 18	19 - 23	24 - 29	30 - 46

Standard Level Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 3	4 - 7	8 - 12	13 - 18	19 - 23	24 - 29	30 - 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 applied a single strategy to all questions, regardless of the command term. This included writing out the aims of the level of analysis, writing out a history of the level of analysis and, in the case of section A questions, needlessly 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, in these cases, only the first example is assessed.

Candidates often focused too much on describing research studies. In the case of the biological level of analysis essay question, candidates often described four or five brain imaging techniques. It is important for candidates to adopt a “less is more” approach in which there is

a clear balance between the critical thinking and the content. Many responses were solely descriptive; using fewer examples effectively to meet the demands of the question would make for stronger responses.

There were many poorly written introductions to essays that simply recounted information but were not focused on the demands of the question. Candidates are penalized for the loss of focus in "organization". When asked for one research method, they should not list all the research methods or describe them all. This was a problem in many of the questions where candidates wrote long, unfocused and often irrelevant introductions that detracted from the overall quality of the response.

Critical thinking was often not relevant to the demands of the question. When asked to discuss research methods, arguing that HM could not give consent is not highly relevant to the demands of the question. In order to earn high marks in critical thinking, it must be relevant to the demands of the question.

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

Overall knowledge of the three levels of analysis was satisfactory with the majority of candidates able to identify appropriate concepts and theoretical explanations and provide relevant research examples.

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

Section A

Biological level of analysis

There were many good responses to this question which clearly outlined the evolutionary explanation and then linked it to a behaviour – for example, disgust, attraction, homosexuality or altruism. Many candidates, however, simply described and evaluated a study without addressing the evolutionary explanation.

In addition, many candidates believed that adaptation to the environment was a choice and did not recognize the role that genetic mutation plays in the process. On the other hand, there were several candidates who wrote about genetic arguments, such as the nature of intelligence, instead of addressing evolutionary explanations.

Cognitive level of analysis

This question asked candidates to consider "why" a research method is used at the cognitive level of analysis, but many candidates simply described a study in great detail or outlined how the research method is applied. Good responses, for example, addressed issues of reliability, internal validity, ability to establish causality, ecological validity and richness of data obtained.

Many candidates also struggled to link the research method to the cognitive level of analysis. Many wrote about localization of function, role of hormones and other biological research that is more appropriate to the biological level of analysis. Although brain imaging techniques were accepted as a research method, very few candidates could actually link it successfully to the study of a cognitive process. It is recommended that candidates not use brain imaging techniques as a research method, as it is more correctly described as a data collection technique.

Sociocultural level of analysis

This question was well answered by many candidates. The theory was clearly outlined and explicitly linked to a well described study. Weaker candidates tended to focus on the description of the study while listing the components of the theory without showing understanding of their meaning.

Section B

Biological level of analysis

There were many very strong essays for this question. The stronger essays focused on one or two brain imaging technologies and discussed their use in detail.

Weaker responses were highly descriptive of the studies, but did not clearly address the brain imaging technologies. There was often superficial evaluation which focused on practicalities – without explaining how those practicalities may have an effect on psychological research; for example, the issue of claustrophobia. This may cause the amygdala to activate and thus influence the observed brain activity in an fMRI. Several limitations were not linked to research. For example, many candidates wrote that participants may not have any metal whatsoever for an MRI. Besides not being true, this was not well linked to the question of the use of technology in the investigation of the relationship between biological factors and behaviour.

Cognitive level of analysis

There were many weak responses to this question. Candidates often overstated claims of research being “unethical” with minimal understanding of the research or how ethical considerations are applied – for example, that consent can be obtained from guardians; or that informed consent explained the procedure so that participants were not surprised, for example, by Sharot asking them about 9/11.

Many candidates did not focus on only two ethical considerations. The strategy to define and discuss all of the ethical considerations in the introduction is a very poor strategy and should be discouraged.

Sociocultural level of analysis

There were many very good responses to this question. Many candidates clearly described research and then linked it to the compliance technique. Strong responses met the demands of the command term and discussed the underlying assumptions and/or theories upon which the compliance techniques are based.

Many responses, however, were solely descriptive or the critical thinking was not relevant to the demands of the question. Simply evaluating the studies is not always the appropriate critical thinking strategy to meet the demands of the question; for example, discussing the sampling techniques and/or ethics of Cialdini's research is not highly relevant to a discussion of underlying theories and principles.

Candidates often outlined the factors that influence compliance techniques – such as liking and scarcity – without making any connection to the two techniques that were examined. This makes the response less focused.

Recommendations and guidance for the teaching of future candidates

Teachers should 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. The command terms should be clearly reviewed by teachers in order for candidates to correctly focus on the requirements of the questions.

Candidates should also be advised that when the question asks for only one study or theory, only the first one addressed by the candidate is assessed. Teachers should apply this method of assessment in their own practice in order for candidates to break the habit of telling all they know, rather than focusing on the question.

More time should be spent on developing critical thinking skills. Many candidates had a very limited range of strategies which they applied without regard to the demands of the question.

Finally, it is important that during revision for exams candidates be guided in their understanding of how to apply studies to address the learning objectives. Several candidates did not use studies that were relevant to the level of analysis stated in the question or they used examples of research which were not good examples – for example, Brewer and Treyens, or Bartlett as examples of experiments. Even though candidates may earn marks if they clearly make those links, the majority of candidates that use these examples are not able to do so.

Higher and standard level paper two

Higher Level Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 5	6 - 11	12 - 16	17 - 21	22 - 26	27 - 31	32 - 44

Standard Level Component grade boundaries

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

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 studied.

Candidates should be continuously reminded that all questions included in paper 2 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 did not focus on what was being asked, but 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 closer to the previous programme 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 skillful ways. 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 a very popular question within the option and probably together with question 2 and question 11 one of the most popular in the whole exam.

The best answers identified a disorder and discussed one model or theory of the disorder. Discussion of the model or theory was usually attempted by providing strengths and limitations of the model or theory and supporting the argument with relevant empirical studies. Unfortunately many responses overly focused on description and explanations of models or theories and made a minimal effort to address the required discussion.

Responses which tended to get fewer marks were those that provided long and descriptive accounts of one disorder with minimal reference to relevant models or theories. Many candidates struggled with the term “theory or study” and wrote everything they knew about a specific abnormal behaviour. Also many weaker responses provided a general response on definitions of normality and abnormality – these answers could only receive marks in the lowest markband since the majority of the response was not addressing the question.

Many candidates decided to discuss one study (rather than theory). The most common responses made reference to Rosenhan’s study and the description and evaluation of the study was often done in a superficial and general manner with many inaccuracies and lacking relevant information (what type of study was it, how were the data collected, what were the actual results and how were they interpreted). Some candidates misinterpreted the question and discussed specific abnormal behaviours (etiology, morbidity rates, etc.)

Question 2

This was also a very popular choice and the quality of responses tended to vary greatly. In the majority of cases, cultural considerations were addressed in a general and vague manner. Candidates referred to the following issues: culture bias, examples of culture-bound syndromes – often vaguely explained, differences between individualistic and collectivistic societies and differences between developed and developing countries.

Instead of addressing the question some candidates chose to discuss how social factors influence concepts of normality and abnormality or described different classificatory systems used in different parts of the world. When a lack of focus was present in the response candidates tended to make reference to Rosenhan’s study and other studies indicating problems with diagnosis but with no relation to cultural consideration.

Higher quality responses tended to discuss the following issues:

- how different cultures define abnormality
- difference in prevalence rates across cultures
- changes in culture over time

- cultural differences in presentation of symptoms
- emic versus etic approaches to diagnosis

Question 3

This was the least popular question. Strong responses chose a specific disorder and then compared and contrasted one biomedical and one individual approach to treatment. Usually drug therapy was contrasted with the cognitive approach. More emphasis was given to biomedical treatment.

In weaker responses, candidates gave a general and vague description of drug treatment and cognitive therapy without addressing the command term "contrast". In these answers candidates described biomedical treatment in detail and provided evaluation; then outlined cognitive therapy and described research studies that claim cognitive therapy is equally successful as drug therapy. Very often only implicit contrast was provided by statements such as "cognitive therapy has no side effects".

A few candidates used group therapy instead of an individual approach to treatment.

Developmental psychology

Question 4

This was not a very popular question, and answers to this question tended to be poor to mediocre. It seemed the phrase "social and/or environmental variables" was problematic. Answers using Piaget did not refer to social/environmental variables at all. Answers using Vygotsky usually described the zone of proximal development and scaffolding, but again failed to highlight the social aspect of these concepts. Better answers used social variables such as low socioeconomic status, low education levels in parents, and malnutrition (as a result of poverty) to support the argument. Some candidates cited studies on feral children and children of extreme neglect (e.g. "Genie"; Koluchova's case study of twins) to support their answers.

Question 5

This was the most popular question within the option. Most responses started off with a clear definition of attachment. Many candidates devoted a large part of their response in defining and describing different types of attachment. Unfortunately, these long descriptions were often not focused or made relevant to the specific question stated.

The term "later in life" was often ignored and candidates discussed immediate reactions of children to separation.

Candidates also failed to provide clear and focused evidence of critical thinking. Some attempts were made to use relevant psychological theories and studies in order to discuss the link between attachment in childhood and the effects this could have on the formation of relationships later in life. However, evaluation of theories or studies was usually provided in a broad and general manner (by referring to methodological, ethical, cultural considerations) rather than addressing the basic question – how do these theories and/or studies explain the link between childhood attachment and the formation of relationships later in life?

Better answers used Hazan and Shaver (1987) as a means of discussing relationships later in life. A few candidates used Erikson's first stage of development in a useful manner. A few candidates mentioned the topic of "resilience", but did not do a good job of relating it to the question.

Question 6

This was the least popular question within the option. This question was not well answered in most cases.

The biggest problem was that some candidates misinterpreted the question. The question was understood as requiring two theories/studies "leading up to adolescence" rather than "into adolescence". Some candidates described the stages of development prior to adolescence (gave an overview of Piaget's, Vygotsky's or Erikson's theory), albeit with a great degree of accuracy, but it was largely irrelevant regarding answering the question being asked.

Health psychology

Question 7

This question was rather popular within the option. The responses tended to provide rather detailed information and contained quite a lot of knowledge and reference made to relevant empirical studies – usually social factors were addressed (social learning theory, peer pressure, influence of media and cultural differences) and biological factors (genetics, dopamine levels).

Overall, mediocre responses lacking explicit evidence of critical thinking relevant for the question were provided. Often empirical evidence was provided but wasn't used effectively for the specific question.

Question 8

This was a rather popular question within the option. Unfortunately a large number of responses provided common sense knowledge about the benefits of exercise, yoga or gastric surgery.

Some higher quality responses presented several studies that compared the effectiveness of different treatments.

Less prepared candidates tended to write about many factors related to overeating and the development of obesity in a broad and superficial manner with minimal reference to treatment.

Question 9

This question was not a very popular choice. Usually when it was addressed it contained relevant knowledge and understanding of the topic but evidence of critical thinking tended to be rather basic. Most candidates chose to write about two or more health promotion strategies reflecting that they probably didn't have a deep understanding of only one health promotion strategy. Popular choices were:

the Victoria (Australia) campaign, “Go for your life” promoting healthy eating and exercise in schools (2004) and the Florida (US) campaign, “TRUTH” an anti-smoking campaign arranged by and aimed at adolescents (1998–1999).

Some candidates misinterpreted the question and wrote about how people might promote healthful living in their own lives. These responses were not well presented and showed a lack of critical thinking as well.

Psychology of human relationships

Question 10

This question was not popular within the option. A number of different cognitive factors were addressed: self-esteem (Kiesler and Baral’s (1970) study of the role of self-esteem in attraction); social learning theory in relation to violence; attribution; and language/communication.

Candidates also referred to a variety of different human relationships including romantic relationships, friendships, social responsibility or violence.

Most responses reflected good knowledge and understanding of how several cognitive factors influence relationships. However, a discussion was usually less thorough.

Candidates who chose a small number of cognitive factors usually provided more clear and detailed evidence of critical thinking.

Question 11

This was the most popular question within the option and one of the most popular questions in this exam. Most candidates tended to explain cross-cultural differences in prosocial behaviour by explaining the relevance of: cultural norms; different socialization processes in an individual’s upbringing; or cultural dimensions (for example, individualism versus collectivism).

Candidates usually made reference to the following studies: Whiting’s (1979) research on the role of extended family; Levine’s studies on cultural differences in prosocial behaviour; or Whiting and Whiting’s (1975) research into altruism levels in children from industrialized and non-industrialized countries.

The majority of responses tended to reflect good knowledge and understanding and at least basic evidence of critical thinking skills. Although cross-cultural differences in prosocial behaviour were usually addressed, candidates failed to provide an explanation of these differences.

Question 12

This was not a very popular question. Most responses tended to refer to the following social origins of attraction: proximity; cultural norms; the mere exposure effect; social exchange theory; or social identity theory.

A wide variety of research studies was used in order to support the information presented. Some popular choices were:

- Simmons et al.'s (1986) study investigating cross-cultural differences in the way romantic love is valued
- Buss's (1994) questionnaires on mate selection
- Zajonc et al.'s (1960s) studies on the mere exposure effect.

Candidates usually provided evidence of critical thinking by addressing alternative factors (cognitive and biological), providing supporting evidence or by addressing methodological considerations. Usually, good responses were provided although one origin of attraction was often more thoroughly discussed than the other.

Sport psychology

Question 13

Candidates responding to this question generally discussed internal versus external motivation without the benefit of reference to studies showing the relative effectiveness of each.

Question 14

Candidates responding to this question generally discussed using mental imagery and self-talk for skill development, although many referenced either personal examples or referred to research in a vague manner.

Question 15

A range of different responses was provided by candidates. In some cases this question generated relevant and focused responses. However, the majority of responses listed reasons for using drugs in sport (including names of a few drugs) but did not support their knowledge with empirical evidence. Again, candidates used anecdotal reports of specific athletes who have been reported to use drugs in sport, particularly in the instances of performance enhancing drugs.

The command term "discuss" was not well responded to and answers tended to be one sided.

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 candidates had problems in structuring a response to "contrast" questions. 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 final 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 - 14	15 - 17	18 - 30

General comments

The stimulus material seemed appropriate and most candidates were able to use it to some effect in their responses. The topic of this paper was paternal education in Sweden and many candidates were able to understand that this study was a preliminary study of a problem that was perhaps not very general worldwide. The stronger responses showed a clear understanding of how to link knowledge of qualitative research methods to the stimulus material but as always there was a tendency in weaker papers to rely on personal opinions or speculations related to the topic instead of knowledge of qualitative research methodology applied to the study. In the weaker responses, candidates tended to analyse the study and suggest how the researcher could improve the design instead of using the stimulus material to

address the methodological aspects of the study as required by the three questions. As in previous sessions it was surprising that many candidates used the term “experiment” for study and “experimenter” for researcher. Since paper three is about qualitative research methods it is encouraged that teachers have their candidates understand that “experiment” is a specific method among a number of research methods and that “experiment” is not a generic term for “study”. It looks rather strange when a candidate writes "this experiment was a narrative interview".

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

Generally, weaker candidates tended to comment on the study in the stimulus material instead of explicitly addressing the questions asked. Weaker candidates overall had a tendency to show limited or mostly generic knowledge and very limited meaningful integration of the stimulus material into their responses. Typically such responses relied on citations from the stimulus material to answer the questions or they treated the stimulus material as a text that they had to analyse and comment on rather than using it as documentation for their comments related to the questions asked. Weaker candidates also seemed to have problems understanding what was required when the command term “evaluate” was used (as in question 3) – partly because they read this as “evaluate the study”. Many candidates in the lower markbands did not seem to know that they should evaluate the narrative interview used by the researchers in the study (that is, make an appraisal by weighing up strengths and limitations). Instead they either offered their opinion on the study or pointed only at strengths of using the particular method.

In question 1 many candidates focused only on a number of ethical considerations, which meant that the candidate could only score up to a maximum of 5 marks as the question asked for “considerations” and not “ethical considerations” in particular. It also appeared that the command term “explain” was not attended to appropriately in weaker responses.

Question 2 seemed difficult for some candidates although there were some really good answers. The command term "to what extent" appeared to be difficult to address effectively for some candidates. Weaker responses just explained what generalization is with reference to populations and had very limited, if any, knowledge of generalization in qualitative research. Many candidates ended up concluding that the results of the study could not be generalized at all for reasons such as the small purposive sample, that only Swedish men participated, that only men participated or that the participants all came from an industrial town. While some of these points could be relevant if properly explained within the framework of knowledge of generalization in qualitative research, this was mostly not done in weaker responses.

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

In question 1 many candidates used the stimulus material well and could integrate their knowledge of qualitative research methodology with relevant parts of the stimulus material; stronger candidates did an excellent job here. Most candidates demonstrated good knowledge of two or more considerations (for example, sampling and ethical considerations) that could be

relevant before conducting the interviews in the study and were able to link these to the study and explain why they were relevant.

In question 2 some candidates seemed very well prepared and had a good understanding of challenges in generalizing results from a study like the one in the stimulus material. Stronger candidates could refer to specific terminology and the assumptions underpinning generalization in qualitative research and used the stimulus material well.

With regard to question 3 and narrative interviews, some candidates were really well prepared. This question had some very strong responses although there were many weak responses demonstrating very limited knowledge of the narrative interview.

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

Question 1

This question was overall well answered with sound knowledge of relevant considerations – both methodological and ethical – explained in the context of the study. There was a tendency to cover this question in breadth rather than depth but some candidates chose to explain two considerations in depth, and both approaches could earn high marks. Candidates chose between a number of considerations mentioned such as sampling; choice of interview method; choice of interviewer; training of interviewer; choosing a way to transcribe the interview; and decision to record the interview. Strong candidates were able to provide good reasons for why certain considerations were relevant in the context of the study, for example explaining the choice of purposive sampling as relevant to this particular topic that had not been researched a lot previously. Most candidates also explained ethical issues with reference to the stimulus material. The most common considerations referred to sampling, choice of interview method as well as informed consent and acceptance from an ethical committee as ethical considerations.

Weaker responses had a tendency to just briefly describe what was done in the study, relying heavily on quotations from the stimulus material and offering no explanation. It was also common in weak responses for the candidate to offer suggestions as to how the researcher could have done a better job in preparing for the interviews. Or candidates just referred to a number of generic ethical considerations with no specific reference to the study.

Question 2

Stronger responses had a good grasp of the specific conditions for generalization in qualitative research as well as an understanding of how this differs from generalization in quantitative research. Many of the stronger responses were able to demonstrate not only sound knowledge including specific terminology for generalization in qualitative research, for example, referring to transferability to other settings with similar problems (or inferential generalization), representational generalization, or theoretical generalization, and such responses could apply this knowledge to the stimulus material. Stronger responses argued that context and sample in the study provided limited possibility for generalization outside similar contexts unless other

studies could confirm the findings. The best responses also observed that this study was an exploratory case study with limited previous research, which overall could perhaps create a platform for theoretical generalization if further research should be conducted.

Weaker responses demonstrated no clear knowledge of generalization in qualitative research but all had some knowledge of statistical generalization. While it is perfectly acceptable to refer to generalization from populations in a question like this it should be noted that the main focus in paper 3 is on qualitative research. The weakest responses focused on factors from the study, for example sex of the researcher and researcher bias because she interviewed men, and claimed that because of such factors the results could not be generalized; or addressed problems in generalization due to lack of use of postmodern transcription.

Question 3

Stronger responses demonstrated sound knowledge of the narrative interview and its use in this specific study, pointing at both strengths and limitations of the method in relation to the study. Such responses often compared narrative interviews to semi-structured interviews and/or focus group interviews and were able to present an argument of why the narrative interview was chosen by the researcher for its specific strengths.

Weaker responses demonstrated very limited accurate knowledge of the narrative interview, if any at all. Some candidates clearly did not know the difference between semi-structured and narrative interviews. Weaker responses also evaluated the interview process itself rather than the use of narrative interview in the study, for example making reference to the interviewer's behaviour and the process of inductive content analysis. Other weak responses simply described the study in the stimulus material and only sporadically made reference to narrative interviewing. It was also common to see weak answers evaluate the study and not the use of narrative interview in the study, which often resulted in only covering some strengths.

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 terms and using their knowledge of qualitative research as well as information from the stimulus material to support their analysis. 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. Furthermore it is recommended to teach candidates not to write introductions in their responses: first of all this is not necessary in a short answer question and secondly, it often results in repetition thus using valuable time and space to write the same thing twice.

Teachers are recommended to focus on training candidates in using the stimulus material properly so that the content of the stimulus material can be related to relevant knowledge of qualitative research methods. Training could also guide candidates to avoid referring to research methods in a generic way – that is, without much reference to the stimulus material.

Some generic knowledge is necessary in answering the questions but if candidates merely describe what they know within a given area and forget to integrate this knowledge with the study in the stimulus material they will score in the lower end of the markband.

Teachers are recommended to use past exam papers and markschemes when preparing candidates for the exam. This will help them to acquire an understanding of how to apply relevant knowledge and understanding of qualitative research methods to the study mentioned in the stimulus material, as well as communicating this correctly. 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 directly linked to the questions asked. Likewise, it is recommended that candidates be exposed to some qualitative studies during the course to give them an opportunity to fully understand the philosophy of qualitative research. It is also recommended that teachers provide opportunities to practise 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 to prepare candidates 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 in their responses. 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.