PSYCHOLOGY TZ1

(IBNA / IBLA)

Overall grade boundaries

Higher level

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 9	10 - 20	21 - 29	30 - 42	43 - 55	56 - 67	68 - 100

Standard level

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 10	11 - 24	25 - 33	34 - 45	46 - 57	58 - 68	69 - 100

Higher level internal assessment

Component grade boundaries

Higher level

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 generally of a good standard although non-experimental studies were also submitted this year. The research topics were as usual related to cognitive psychology and this approach is more likely to provide good results at this level of education. The majority of candidates were aware of ethical issues and most included a copy of informed consent in the appendices but a few candidates did not obtain parental consent for participants under 16 years.

In general, the weaker reports shared the following characteristics:

- Weak and imprecise explanation of background research in the introduction and this affected the discussion section as well.
- Results were not always clearly related to the aim of the study and inferential tests were absent or not justified.
- Discussions were superficial and did not discuss own results in the light of the background research.
- Referencing was poor.

Candidate performance against each criterion

Most candidates seemed familiar with the assessment criteria although there were differences in achievement levels. The introduction sections were in some cases very well written with a clear focus but it seems that it is difficult for some candidates to have an exclusive focus on the relationship between a particular research study and the candidate's own research hypothesis. The introduction is important in that it presents the rationale for the candidate's own experiment and uses the background research to justify the candidate's own research hypotheses. Therefore the background research should be explained and analysed in sufficient depth (for example, aim, procedure, findings) to allow for formulation of a clear research hypothesis in the introduction and to stimulate discussion of own results in the light of the background research in the discussion section. The experiment should be simple and therefore it is ok to make a partial replication of a research studies, e.g. reducing the number of variables. The introduction and the discussion sections are often the most difficult to write as they require a good understanding of how research studies are linked to formulation of new hypotheses. The level of depth of the analysis of the background research was at times shallow. This influenced the discussion where comparisons of the candidate's own results and those of the background studies could not be done in sufficient depth. Referencing was not always of a standard format and there are still problems with resources found on the internet.

Criterion A Introduction

The aim of the study was not clearly formulated in all reports. The analysis of background research was well done overall but also superficial at times and often based on a summary of a study found on a website or a study guide. This gave some problems in terms of justifying and formulating clearly operationalized hypotheses. The explanations of the theoretical framework and the studies in the introduction lacked depth in some reports, which could explain why it was not always clear from the background readings in the introduction why a particular experimental hypothesis was chosen; however, a considerable number of candidates did a good job.

For some candidates, it was difficult to state a clear and justified experimental hypothesis. Quite a few introductions included redundant explanations of research that was not particularly relevant to the candidate's own study.

It is important that the introduction is clear and focused on relevant background research so that it logically leads towards the candidate's own research hypotheses and that the background research is explained and analysed in sufficient depth (for example, aim, procedure, findings) to allow for formulation of a clear research hypothesis in the introduction and to stimulate discussion of own results in the light of the background research in the discussion section. It should be noted that it is not required to make an exact replication of an experiment. A partial replication will do but the candidate's experiment should be closely linked to an actual experiment and not just inspired by some kind of research. It was obvious that some candidates had chosen this approach and therefore wrote very weak introductions.

Criterion B Method: Design

Most candidates seemed aware of what is meant by an experimental design but choice of design was not always properly justified, e.g. by reference to strengths and limitations of respective designs (repeated measures and independent designs).

The ethical guidelines were mostly addressed in the design section but sometimes in procedures instead and that is also acceptable.

What is essential is that ethical procedures are addressed appropriately somewhere in the report. Most candidates included a blank copy of the informed consent form from participants. A number of schools had allowed participants under the age of 16 to participate without parental consent and this is violating the ethical guidelines of the IB. There were only a few cases where the informed consent form was not included at all and this was always in weak reports where other things were missing as well. As in previous years there were candidates who used many experimental conditions but it is recommended to use two conditions only.

Criterion C Method: Participants

There were often problems in identifying relevant characteristics of the participants. It may be difficult to define relevant characteristics in relation to a specific study but candidates could focus on characteristics such as age, sex, number of participants and nationality. The target population was not always identified. All samples were based on a student population in the candidate's own school. Sampling technique was mostly clearly identified as opportunity sample or self-selected sample and it was often explained or justified. Both ways are fine. Some candidates claimed to have used a random sample, which it was clearly not. There seems to be confusion about the meaning of the word 'random'. Some candidates did not specify how they allocated participants to the experimental conditions but they cannot be penalized for this since it is not mentioned in the assessment criteria.

Criterion D Method: Procedure

There were quite often problems with the description of the procedure in sufficient detail, which makes a procedure difficult to replicate. There was not always reference to the material in the appendices (either here in procedure or in materials) and in these cases replication is simply not possible. It is recommended to include all material and give details on how the material was used.

Criterion E Results: Descriptive

Most candidates described the results in a narrative form in the results section. Not all included standard deviation as descriptive statistics even if their data allowed it. There is a tendency to include a whole range of measures of central tendency in the results section but this is redundant. It does not affect the marks but it is a clear indication that candidates have not considered which measure would be most appropriate for their data.

In some reports there were no graphs but only a table of results. The graphs were often poorly labelled. Not all reports included tables. A few candidates had individual scores in the results section but most candidates had correctly placed the raw data in the appendices. Quite a few candidates graphed their standard deviation side by side with the mean but this does not really make sense and should be avoided.

Criterion F Results: Inferential

Most candidates chose and justified the inferential statistical test correctly but a few did not. Most candidates use the non-parametric tests but an increasing number of candidates use the t-test, which is also appropriate since these tests are quite robust. Not all candidates could justify their choice of statistical test and a number of candidates did not make a statement of statistical significance.

Criterion G Discussion

A number of high-scoring candidates demonstrated competence in discussing own results in the light of the background research and these candidates identified relevant methodological limitations and suggested relevant modifications. However, many were not able to integrate background research properly in the discussion of their own results. Many reports had a very short and superficial discussion of own results in the light of previous research but a long description of limitations of own design, which were not always relevant in terms of their own study but rather a more general list. Some candidates included strengths of own design although this is no longer necessary. The suggestions for modifications were often related to general factors rather than the candidate's own study. For example, there were a lot of candidates who mentioned the opportunity sample as a problem and suggested a random sample and more participants in future research.

Criterion H Citation of sources

The most common problem concerned referencing. Candidates often did not include all the references they mentioned in the introduction. There are still problems with references from the internet where candidates tend to think that the URL is enough and sometimes the background study could not be found in the reference section. Referencing often did not follow a standard way of referencing such as APA.

Criterion I Report format

Some candidates lost marks here because they were missing part of the abstract or missing parts of the appendices.

Recommendations for the teaching of future candidates

- Most important is to help candidates find appropriate background research and theoretical framework because such resources will enable them to analyse the background research in some depth. It is recommended to find relatively simple experiments to replicate.
- This would help candidates to make relatively simple experiments themselves based on real scientific experiments. It is much easier for candidates to replicate an experiment if they have access to readings about experiments or summaries of them so that they can read about hypotheses and other important details.
- There should be more focus on the relationship between the aim of the candidate's study and the background research so that these can be integrated in the introduction and the discussion of the results. The background research should be analysed in sufficient depth in the introduction so that the aim of the candidate's own research is clearly justified and the experimental hypothesis should be clearly linked to background research.
- The design question must have a clear description of the experimental conditions and teachers should ensure that there are only two conditions (either two treatment conditions or one treatment condition and one control) so that there is a possibility to compare the outcome of the manipulation of the IV on the DV in the two conditions. This is in line with the IB recommendations of making simple experimental studies with psychology candidates.

- Sampling should be done according to IB rules, i.e. identification of target population including relevant characteristics and description of sampling method as well as explanation (or justification) of the use of the chosen method. Most candidates use a convenience sample but they should still explain the sampling method. The number of participants in the experiment does not need to exceed 20 (independent design) or 10 (repeated measures design).
- In the descriptive statistics section, the graphs and tables should have a proper title. It should be emphasized that graphing the results is mandatory and that a table must be included. This section includes *summarized* data not raw data or individual scores. It is recommended that candidates don't include several measures of central tendency but only the one, which is relevant for their data, and that a measure of dispersion is present.
- In the inferential statistics section, candidates should be careful in choosing an appropriate statistical test and justify why this test was chosen. This could relate to the level of measurement of data.
- The explanation of the empirical studies and theoretical framework 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 making 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. Understanding of own 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 better. The limitations should be explicitly relevant to the candidate's own experiment.
- It is generally recommended that candidates are familiar with the scientific research method, which includes references to previous studies and integration of these in their own research. Likewise, the use of proper background reading must 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. Furthermore, the candidates could gain more marks in criterion G if they learned to use a standard way of referencing.

Standard level internal assessment

Component grade boundaries

Standard level

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 range of topics were investigated for the internal assessment component for SL psychology. The most popular topics came from cognitive psychology topics as these studies can usually be successfully conducted within a school setting.

A few correlational studies were submitted as well as some quasi-experiments investigating non-manipulated variables (such as gender, age, handedness); both of these types of studies are not acceptable for IB psychology internal assessment.

Many examiners reported there was an increase in the overall standard of the reports. Also, reports from the majority of schools were well marked by teachers. Procedural aspects of the projects were generally well done; however, justification and explanation were weaker. Additionally, the choice and explanation of descriptive statistics posed a problem for many candidates.

Some candidates submitted reports on studies that do not follow ethical standards that have been produced for IB psychology. Teachers should be reminded that candidates in this course have not had sufficient training to appropriately handle topics that may be socially sensitive, overly deceptive, raise anxiety in the participant or that may reinforce negative stereotypical behaviour (especially with respect to gender, race or culture). Therefore such studies should not be conducted as part of this course. It is the teacher's responsibility to ensure that all candidates follow ethical guidelines. The psychology guide (first examinations 2011) has some ethical guidelines for internal assessment that can help teachers handle choice of topics and approval processes. If teachers have additional questions or issues they should seek assistance from the Online Curriculum Centre.

Candidate performance against each criterion

Criterion A Introduction

Introductions had clear aims and relevant research explained in a clear manner at the top of the range. However, many candidates did not explain all relevant aspects of the original study clearly. Often, candidates tended to overemphasize procedural aspects of the study while clearly omitting relevant information about findings. Also, although most candidates presented an aim many failed to state the aim clearly by indicating the IV and DV.

Some examiners have reported that they have occasionally read introductions that are presented as reviews of psychological research – providing description of more than one study and often accompanied with theoretical background. Often, these introductions tended to be long and this clearly reduced the available word count for other parts of the report.

Criterion B Method: Design

Although some weaker candidates still have problems in understanding the difference between a method (experiment) and design (repeated measures, independent samples, matched pairs design) most candidates successfully and clearly identified their research design, IV and DV. Operationalization of these two variables was not always clear. Also, not many candidates appropriately and clearly justified their design.

Ethical guidelines were usually clearly followed and evidence of this was provided within the report. Only a few candidates did not attach a blank copy of the consent form and debriefing letter in the appendices. Candidates from some schools are still unaware that parental consent is also needed where participants are under 16 years of age.

Criterion C Method: Participants

Most candidates chose students from their classes as a sample. Most identified the sampling technique correctly but only rarely justified the chosen method of sampling. The term "random" is still not understood by many candidates, leading to confusing accounts of participant selection – some candidates tend to frequently describe opportunity or convenience sampling as a random sample.

Some candidates had quite large samples in their study; it is recommended that 15 to 20 participants is an adequate size. While large sample sizes are used in academic research smaller numbers are appropriate for the IB psychology internal assessment. The purpose of this coursework is for candidates to get experience in experimental design, not to generate new theory or challenge existing theory.

Criterion D Method: Procedure

In the majority of cases procedure sections were well presented. However, some candidates did not include all relevant materials which they used when conducting their study (standardized instructions, informed consents, lists of words, and debriefing notes were not always included in the appendices). In addition, some candidates included material in their appendices but only for one condition (instead of both conditions) or provided material photocopied in black and white although colour was an important factor – this meant that replication would be difficult.

Criterion E Results

In many cases this was the weakest part of the report. Many problems were noticed by examiners:

- Many candidates presented only measures of central tendency.
- Some candidates presented all measures of central tendency and all measures of dispersion.
- Many candidates presented their data without any explanation.
- Graphs and tables were rarely clearly presented due to inexact or incomplete labelling.
- Weaker candidates chose the wrong type of graph (histograms or pie charts are not appropriate for presenting differences between results of independent groups).
- Some candidates included in this section raw data or graphs showing each individual participant's score.

- Raw data and calculations were often not included in the appendices.
- Calculations (e.g. of standard deviation) were sometimes incorrect.
- Some candidates did not present results that were relevant to the aim of the study.

Criterion F Discussion

Quality of discussions varied but there was a general consensus among examiners that this section was often much better written than in the past sessions. Many candidates discussed their results fully and in a more sophisticated manner.

However, candidates who hadn't clearly explained the study being replicated in the introduction understandingly tended to have difficulty with the discussion.

Another common problem was that often some relevant information was provided in the discussion section but it wasn't fully developed by providing a clear link between the methodology used in the study and the results obtained.

Also, some candidates suggested modifications which aren't really relevant (e.g. doing the study again or having more participants in the study is not a valid modification).

Criterion G Presentation

Reports were generally within the word limit and in the required format. Appendices were appropriately labelled and were referenced in the report. Many candidates are experiencing problems with referencing internet sources and writing clear and concise abstracts.

Some examiners reported reading very low quality reports which were presented in a careless manner and seemed to be finalized in a great hurry. These reports had a number of problems including omitting whole sections of the report, not including all the material used in the appendices, spelling and errors in simple grammar or poor quality of printing and layout.

Recommendations for the teaching of future candidates

In the introduction more focus should be provided on the study replicated. The following information should be included: aim, some details of sample, procedure and findings of the background study.

Ethical considerations need to be clearly addressed in the IA report. This is clearly reflected in the following assessment criteria:

- The assessment criteria for the design section ask for "clear indication and documentation of how ethical guidelines were followed".
- The procedure section asks that "details of how the ethical guidelines were applied are included".

Candidates should document how ethical guidelines were considered and applied in the experiment by including:

- A copy of the informed consent in the appendices
- A copy of the debriefing note in the appendices.

Candidates should also provide a short but clear explanation of how ethical procedures were applied in the experiment – this information can be presented in either the method section or procedure.

Teachers should remind candidates that participants under the age of 16 must have parental permission and this should be stated in the report and documented in the appendix.

When designing the study candidates should select one independent and one dependent variable. There is no need to have more than one IV and one DV – overly complex designs bring complications when analysing and presenting results and lead to very limited and superficial discussions.

Candidates should report relevant characteristics of participants – these could be age range, sex, nationality, psychology students vs. not psychology students or characteristics that are important in the specific experiment (such as not suffering from colour blindness in a Stroop experiment).

Candidates should develop skills of interpreting results and analysing the data that is collected. In the results section, candidates should ensure they provide table and graph headings and provide sufficient description of what they reflect. It is important that candidates specifically name their average scores; do these reflect mean, median, mode? Also, candidates should describe what these differences between scores of experimental and control groups reflect; and importantly what the SD or range imply. Candidates should be encouraged to check all calculations, clearly write them and include this in the appendices.

More effort needs to be given in to making sure that candidates understand the expectations of the IA and the specifics of the assessment criteria descriptors. Currently many candidates are missing marks simply because they are ignoring some of the required elements.

Some SL candidates write their reports by including elements required from HL IA (e.g. writing a hypothesis or applying inferential statistics to results). Candidates should be clearly warned not to do this because only SL assessment criteria can be applied to all SL reports. Therefore all unnecessary information will be ignored but included in the word count.

Abstracts should be written with care. The format of the abstract should include the aim, relevant aspects of the procedure, main findings and a short conclusion.

Higher and standard level paper one

Component grade boundaries

Higher level

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 5	6 - 11	12 - 14	15 - 19	20 - 25	26 - 30	31 - 46

Standard level

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 5	6 - 11	12 - 14	15 - 19	20 - 25	26 - 30	31 - 46

General comments

Candidates generally demonstrated adequate content knowledge but had difficulty giving answers focused on the question asked. It was obvious that the skill to structure and develop a clear and coherent argument was a weakness of many candidates, especially in section B. Arguments were either too superficially developed or had not been structured well so that the essay did not flow clearly. Few essays contained a clear introduction to the question and key concepts were only occasionally defined. In many cases the conclusion was very superficial and candidates did not link back to the question or the command term so that the response could not be seen to focus fully on what the question required. Evaluation was also rarely well integrated throughout the essay and in quite a few cases was just added as an afterthought.

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

Understanding the demands of the question continues to be an issue with many candidates. They do not answer the questions accurately and provide marginally relevant information instead. Question 4 in particular seemed to present a great deal of difficulty and few candidates were able to address genetic research in any depth with regard to ethics.

The majority of candidates were able to provide descriptive responses for many questions, but few were able to offer a relevant critical evaluation. Candidates seem to be over prepared in giving formulaic evaluations and criticisms of studies. Often these were written after each study even though they were irrelevant to the question. Studies were criticized for being “ecologically” valid (or not) with no attempt to explain what this means or for having small samples or a lack of “other” cultures with no attempt to justify these claims.

The areas of the programme and examination in which the 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.

Many candidates answered well questions on hormone, schema theory and conformity. A good number of candidates were skilled in presenting research studies and knew such studies well and precisely.

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

Section A

Biological level of analysis

Many candidates were able to correctly identify a hormone and then provided an explanation of its effect on behaviour. Some candidates identified a neurotransmitter; this received no marks. Several candidates failed to identify the behaviour that was related to the hormone; instead they often described physiological changes. Many candidates used Schachter & Singer's study on cognitive appraisal or Sapolski's study on how social hierarchy among baboons affects stress levels but failed to effectively link those studies to the requirements of the question.

To earn top marks it was essential that the effect be explained. Good responses, for example, addressed the permissive effect of testosterone on aggression, the potential long-term effect of cortisol on memory, or the effect of melatonin on sleep patterns and offered relevant studies to support their explanation.

Cognitive level of analysis

Most candidates could provide an example of a relevant study but it was not always clearly linked to schema theory. To earn top marks, responses should clearly relate the conclusion of the study to the theory. A large number of candidates were familiar with Bartlett's, Anderson & Pichert, and Loftus' studies involving schema in memory but never related schema theory to them. Many candidates tried to rationalize that a study such as Kohler's, Tolman's, Bandura's, or Seligman's, was related to schema theory when it had definitely not been conducted as an investigation into this theory. A minority of candidates fully described the study, focusing on the aim, method and findings.

Sociocultural level of analysis

Some candidates offered well developed responses for this question describing fundamental attribution error with reference to the Ross et al. study and self-serving bias with reference to the Kashima & Triandis study. However, there was often confusion between the various errors and some candidates discussed cognitive "errors" which were not clearly attributional ones. Some candidates wrote about "sociolearning perspective" and described Watson's, Skinner's or Kohler's studies. Unfortunately, these candidates seemed to still write about the learning perspective included in the previous programme.

Section B

Biological level of analysis

This question was rarely answered. In the majority of cases it was not well answered as candidates found it difficult to address genetic research. A number of unrelated theories and studies were used such as brain localization and Gage case study.

There were very few candidates who identified more than one relevant example of research and for the most part these essays examined very general ethical considerations and failed to make a meaningful link to concerns particular to genetic research. Some candidates appeared unsure as to what genetic influences on behaviour meant; there was confusion between genetic and biological influences. Other candidates focused on the nature vs. nurture debate. For the most part, the command term "discuss" was not addressed and critical thinking relevant to the question was not offered.

Cognitive level of analysis

Most candidates focused on memory as the cognitive process and were able to support their answer with relevant studies such as Ross & Milkson, Martin & Halverson, or Cole & Scribner. Cultural and/or social factors such as education, gender differences or memory strategies in different cultures were addressed and this was achieved quite competently. Critical thinking was often limited though and few candidates gained marks in the higher markbands for criterion B because answers were descriptive rather than discursive and analytical. The "extent to which" command term was not well focused on for the most part and conclusions tented to be very general. Focused answers that did present relevant critical thinking discussed how different factors, in addition to the cultural and social ones, might affect memory such as emotions in flashbulb memories or biological factors in amnesia. Some candidates failed to identify a specific cognitive process, writing only in general terms.

Sociocultural level of analysis

Many candidates wrote a very well developed discussion of the factors that influence conformity; however, there were several candidates who simply outlined the nature of conformity and identified factors that would increase or decrease the likelihood of conformity to the group, but offered a very limited discussion. Factors included the size of the group, cultural dimensions, gender, anonymity, self-esteem, minority opinion and the phenomena of group-think and risky shift. Strong responses were able to discuss several factors in depth, providing relevant research to support their answer.

Weaker candidates discussed obedience or compliance techniques rather than conformity. Many candidates who described the Stanford prison study were unable to identify the factors that were responsible for the observed level of conformity.

Recommendations and guidance for the teaching of future candidates

Teachers need to focus on developing essay writing skills. Candidates need to be shown how to go beyond a descriptive response and that knowledge has to be evaluated and analysed for higher marks to be gained. A clear definition, a good example and a piece of research highlighting the phenomenon in question plus a clear and coherent argument will gain high marks. Candidates should avoid general and vague statements, lack of precision, use of anecdotal evidence and sensational stories.

Evaluation is also an area where candidates need to be supported by teachers as it was rarely well integrated throughout the essay. Evaluation of studies alone is not enough for the highest marks in criterion B. Candidates should be advised that critical analysis and evaluation has to be related to the requirements of the question. Candidates need to go beyond description and evaluation of studies and focus more on how they answer a question through the analysis, application and evaluation of the findings from them.

Clarity of expression is another issue that teachers should advise candidates on. Candidates should define terms and be sure that theories and studies are clearly explained, not assuming too much knowledge on the behalf of the reader.

Candidates should also be aware that more is not necessarily better and avoid writing long general introductions, detailed description of research or evaluation when it is not required.

Candidates should also be advised to read the questions carefully and give answers on the question asked. When it says “one” it is not to their advantage to write several examples. This is often done to the detriment of the overall quality of the response.

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 - 13	14 - 19	20 - 25	26 - 31	32 - 44

Standard level

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

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

Answers scoring in the lower ranges had obvious difficulties in structuring a response – poor organizational skills, a tendency toward anecdotal comments or generalized responses lacking in specifics. In addition, many candidates found it difficult to support ideas with relevant psychological research so this is an area that could be improved upon.

Candidates had difficulty with questions using “compare and contrast” as the command term. Many responses would contain two separate descriptions of theories, models, strategies, etc. with a weak paragraph at the end in which a few differences were mentioned. Candidates are better at identifying differences than they are at identifying similarities.

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

Overall, candidates were far better able to understand the demands of the questions this year. Many examiners reported noticing a marked improvement over last year.

Many schools 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. Cultural 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.

While there was a range in performance, some candidates were very well prepared in the human relations option, demonstrating a good understanding of basic concepts and relevant studies.

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

Abnormal psychology was the most popular option. Other very popular options were health psychology and psychology of human relationships.

Abnormal psychology

Question 1

Probably the most popular question. Popular choices of disorders were depression, phobia, anorexia and bulimia. A few candidates chose schizophrenia and unfortunately lost all marks since schizophrenia does not belong to these groups of disorders.

More focus was provided for biological factors (detailed descriptions of hormones, neurotransmitters, drugs that help, parts of brain responsible for disorders; research studies and theories offered) than for sociocultural factors (often casual, common sense answers were provided – better answers provided information about cultural differences or sometimes discussed the effects of gender roles; some responses provided research studies on how life events trigger a certain disorder). Answers about eating disorders almost always used the media as a sociocultural factor. Answers in which the diathesis-stress model was discussed tended to be strong responses.

Overall mediocre to good answers. Although several responses were simply descriptive, others included pointed analysis and evaluation in a skillful manner.

Question 2

Ethical considerations were usually more fully addressed than cultural factors. Issues discussed were: stigmatization, labelling, self-fulfilling prophecy, diagnosis leading to treatment which may have negative 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) where viewers were shown a tape of an individual telling of his job experience. Vague and inaccurate descriptions were provided.

Cultural considerations usually were addressed in a general and vague manner. Candidates referred to the following issues: cultural bias, culture bound syndromes, cultural bias and cultural ignorance on the part of the diagnostician, and cultural bias of classificatory systems.

The biggest problem was that instead of addressing the question some candidates chose to discuss concepts of normality and abnormality or described classificatory systems in a detailed manner without focusing on diagnosis. Overall good general knowledge but not many excellent answers.

Question 3

This question was the least popular one in this option. Strong responses chose a specific disorder and then compared and contrasted one biomedical and one individual approach to treatment. Usually drug therapy was compared to the cognitive approach. More emphasis was given to biomedical treatment. Similarities were not well presented. It was easier for candidates to focus on differences.

The biggest problem was that candidates gave a general and vague description of drug treatment and cognitive therapy without comparison. 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 as 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 as an individual approach to treatment. With some exceptions, many candidates had problems comparing and contrasting and instead provided detailed description and evaluation of two theories.

Developmental psychology

Question 4

Answers to this question were uniformly poor to mediocre. It seemed the phrase “social variables” was problematic. Answers using Piaget did not refer to “social variables” at all. Answers using Vygotsky usually described the ZPD 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”) to support their answers.

Question 5

Attachment was usually reasonably defined. Although the command term “define” is very basic (it is associated with assessment objective one indicating that this part of the question has a low cognitive demand) candidates devoted a large part of their response in defining and describing different types of attachment. Unfortunately, these long descriptions were often not given credit as the information provided was not focused on the specific question stated. For

the second part of the question the term “later in life” was often ignored and candidates discussed immediate reactions of children to separation, e.g. Ainsworth’s studies.

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 formation of attachments in childhood and effects this could have on the type 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 & 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 number of candidates mentioned the topic of “resilience”, but did not do a good job of relating it to the question.

Question 6

Very often this question was answered in a broad and general manner by providing a few theories and/or studies on the development of gender roles and providing some evaluation of these theories. In many cases, responses to this specific question reflected that candidates had general knowledge of theories of development of gender roles. However providing a range of arguments and factors about how and why gender roles develop was much more challenging.

Better answers discussed biological factors such as differences in gender behaviour produced by evolution, the role of hormones in the development of gender roles, and then discussed sociocultural factors such as cultural norms, social learning theory, and social role theory in the development of gender roles. Some candidates made good use of the David Reimer botched circumcision case study to show the influence of biological factors in gender role development and Margaret Mead’s research to show the influence of sociocultural factors in gender role development. Strong responses argued that both biological and sociocultural factors were important in gender role development.

Health psychology

Question 7

This question was a popular one within the option. Some responses reflected good knowledge and understanding of strategies for coping with stress, making reference to models such as Lazarus and Folkman's model, health realization model and social support networks. A substantial number of responses gave anecdotal and vague answers in which exercise, meditation, healthy lifestyle etc. were discussed without any use of studies for evaluation. Responses addressing ineffective and unhealthy coping strategies very often provided superficial information lacking in-depth understanding of the topic.

Question 8

This question also was a popular one within the option. Most responses provided a detailed description of relevant factors using clear and precise terminology. However, the next step which should have included a discussion of these factors usually was less thorough. Higher quality responses successfully addressed the command term “discuss” by providing clear

arguments and supporting these claims with empirical evidence, theories or possible applications.

Stronger candidates usually decided to choose only two factors and provided an in-depth discussion of how these factors were related to overeating and the development of obesity.

Most candidates discussed biological and sociocultural factors (usually more emphasis was given to biological factors). 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 psychological research.

Question 9

This question was not a very popular choice. Unfortunately, some weaker candidates ignored the last part of the question and provided a general evaluation of a model/theory related to health psychology in general. These responses did not attract many marks as they are of marginal relevance to the specific question stated. Higher quality responses made an in-depth appraisal of one model – this was achieved in a number of different ways but most candidates decided to compare one model with alternative models of health promotion or provided empirical evidence supporting or contradicting the chosen model.

Psychology of human relationships

Question 10

Responses to this question were generally poor. A number of responses were not informed at all by psychological research. Some candidates turned “violence” into “prejudice” and spent the rest of their time addressing this topic. Some of the better responses to this question focused on “bullying” as the variety of violence addressed and discussed anger management and “whole school” programs. It was a rare response that addressed the issue of the extent to which a violence reduction strategy is effective in any competent manner.

Question 11

This question was very popular within the option. A number of different biological factors were addressed: neurotransmitters, hormones, “pleasure centres” in the brain, centres of “aggression” in the brain, evolutionary theory applied to natural selection survival traits and sexual selection applied to mate selection.

Some candidates interpreted the phrase “human relationships” as “interpersonal relationships” and therefore focused only on attraction and mate selection, but still produced good responses. Other candidates focused on all three parts of the option: social responsibility, interpersonal relationships and violence. This approach also produced some very good responses. The use of the command term “discuss” that many candidates can usually handle reasonably well and the single focus on concrete biological factors combined to produce a significantly higher percentage of good and very good answers than usually exhibited in response to a question.

Question 12

Many candidates did a good job of identifying and explaining two theories of altruism. The difficulties arose in connection with the command term “compare and contrast”. Many candidates wrote separate descriptions and internal evaluations of two theories but provided little or no discussion of their similarities and differences.

Some candidates mixed parts of one theory with another theory leading to a poor answer. The best responses stemmed from the candidate choosing two theories clearly distinct from each other, e.g. kin selection/selfish gene vs. empathetic concern which made contrasting the two theories relatively easy, leaving the candidate to make a point or two about how the theories were similar, e.g. both aimed at explaining altruistic behaviour, both based on theory/studies, etc.

Sport psychology

Question 13

This question was the least popular choice within the option. In general, responses provided rather good descriptive knowledge which tended to lack reference to theories and/or studies.

At times this question attracted the attention of less prepared candidates who provided general information in a superficial way.

Question 14

This question was the most popular one within the option. Many responses reflected detailed knowledge and understanding of theories. Most popular choices were the following theories:

Yerkes and Dodson (1908) inverted-U theory, Hanin's (1997) optimum arousal theory, and Bandura's (1991) self-efficacy theory.

Higher quality responses clearly addressed the command term "discuss". This was done in several different ways: by discussing strengths and limitations of theories or by focusing on empirical evidence that supports or contradicts the theory. In the majority of responses there was more focus on detailed description than on providing evidence of critical thinking.

Question 15

This question was also a rather popular choice within the option. At times answers lacked reference to psychological research and responses provided detailed but superficial accounts of the role of the coach on team behaviour and performance.

Two types of problems were recognized in responses to this question:

- The "to what extent" command term was either misunderstood or completely ignored. These responses provided good descriptions of ways in which the coach affects team behaviour but provided no evidence of critical thinking.
- Although some candidates provided psychological research in their response the research was either not specifically related to the effect the coach has on team behaviour or the evaluation of the study was provided in a general manner and therefore not really linked to the demands of the question.

Recommendations for the teaching of future candidates

- Teaching candidates how to construct an organized response is a very difficult task and therefore enough time and effort should be provided to make sure that all candidates understand how to approach questions and how to structure their response.

- Teachers should encourage the use of terminology relevant to psychology. Many examiners are complaining responses are too general and lack 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 examination paper. Candidates should continuously be reminded to support their arguments with relevant psychological theories/studies.
- If a command term is associated with assessment objective one this indicates that this part of the question has a low cognitive demand and therefore candidates should not devote a lot of time or space to answering this part of the question.
- Several examiners reported that it appeared that candidates had problems in structuring a response to “compare and contrast” questions. Teachers should emphasize the need to satisfy the demands of the specific command term used in the question. For example, when “compare and contrast” is used, candidates must discuss similarities and differences. Examiners also report difficulties with the command term “evaluate” which requires candidates to discuss strengths and limitations. Time spent on making sure that candidates understand the specific demands of the command terms will be time well spent in improving candidate performance in examinations.

Higher level paper three

Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 2	3 - 5	6 - 7	8 - 11	12 - 14	15 - 18	19 - 30

General comment

There was an extremely wide range of marks for this paper that varied from those candidates who had prepared themselves in a diligent manner for the examination to those who appeared to have made little effort to obtain relevant knowledge or understanding of the syllabus requirements.

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

The most difficult question for some candidates involved a knowledge and understanding of inductive content analysis. While some candidates scored high marks for this question there were others who simply did not appear to be familiar with the phrase ‘inductive content analysis’, despite the fact that it is clearly included on the syllabus.

It is generally accepted that work at diploma level should go well beyond the memorizing and subsequent regurgitation of terms. The understanding and application of subject terms is a requirement that candidates need to practise and be able to use under examination conditions. Although a few marks were allocated for mainly descriptive responses these had the potential to be much higher had they been applied explicitly to the stimulus material.

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

In their answers many candidates showed an adequate or higher knowledge of the key terms included in the questions, including those such as participant expectations, semi-structured interviews and inductive content analysis. Regrettably several failed to capitalize on their knowledge when they answered questions in a most general way by making no reference to the stimulus material.

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

Question 1

The term 'participant expectations' has been interpreted in various ways by different researchers and authors. Qualitative research, by its nature, tends to be more fluid than its quantitative counterpart. Provided that an answer contained a reasonable understanding of participant expectations it was given credit. Candidates achieved higher marks if, in addition, they incorporated relevant examples from the stimulus material to illustrate their explanations.

For example the befrienders may have anticipated approbation from the befriendees for what they regarded as their own altruistic behaviour. This interpretation could well be erroneous and lead to incorrect findings by the researchers. Several other examples from the study could have been used.

Question 2

Semi-structured interview techniques were well known to the extent that several candidates were lured into providing lengthy descriptive and general accounts of this method, which were of marginal relevance unless the descriptions were explicitly linked to the study. This occurred despite the brevity of the question that required candidates to 'Discuss the use of semi-structured interviews in this study.' Most candidates knew that richer data can be obtained as a result of the flexibility that this type of interview encourages; it allows respondents to think outside of the regimented and constraining list of questions used by structured interviews.

But this knowledge needed to be put into context so that, for example, befrienders were able to express their appreciation for being able to learn about different cultures, or to empathize with those who suffer social difficulties as a result of coming from a minority ethnic group.

Question 3

Inductive content analysis was a less daunting aspect of qualitative research than some candidates might have anticipated. There were different ways of using such a method. Provided that the interpretation of induction was both clear and relevant, and the fact that the transcript was analysed in a systematic manner, credit was given for such information.

Credit was also given for answers that referred to potential dangers or advantages in the use of inductive content analysis, providing that these included reference to the stimulus material. For example this could have mentioned the danger that some researchers may have been biased for or against the benefits of a befriending scheme, and that unless measures were taken to ameliorate such bias that the findings of the study would contain serious errors.

The type of assistance and guidance that teachers should provide for future candidates

A major difference between the present syllabus for paper three and its predecessor is in the use of stimulus material on which the questions are based. The aim is to base questions on problems that can and do arise during research.

It is recommended that teachers of paper three should provide their candidates with stimulus material that is either taken from published psychological studies or created by the teacher. Questions can be created by inclusion of terms that are shown in the syllabus. The question paper should make it clear that questions should be answered in the context of the stimulus material.

The May 2011 examination has stimulus material in which every fifth line is numbered. This is so that candidates may refer to the lines without having to use extensive quotations.