PSYCHOLOGY TZ2

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

Overall grade boundaries

Higher level

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 7	8 - 17	18 - 27	28 - 39	40 - 52	53 - 64	65 - 100

Standard level

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 10	11 - 20	21 - 31	32 - 43	44 - 55	56 - 67	68 - 100

Higher level internal assessment

Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 2	3 - 5	6 - 8	9 - 11	12 - 15	16 - 18	19 - 25

The range and suitability of the work submitted

A wide variety of work was submitted. Almost all work was targeted at appropriate topics and constructed with appropriate designs. The most popular area of research was from the cognitive perspective, while a few were also from social psychology.

Most schools chose suitable research topics that required an appropriate manipulation of an IV and a DV. Some schools allowed students to conduct experiments on conformity, which is not allowed.

There were a few studies that failed to manipulate an independent variable. For example there were studies that used left- and right-handedness 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 in order to qualify as an experiment for the purposes of the IB psychology internal assessment.

There were quite a few papers that qualified as non-experimental as they were not in line with the definition of an experiment in the marking instructions. 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. Many serial position effect studies were surveys as the independent variable was not manipulated.

There were a small number of sample reports in which students did not clearly understand the experimental nature of the intended study. A slightly larger minority of 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 in depth analysis of previous research and theory that was relevant to the topic. Other papers clearly explained a relevant theoretical model and analyzed previous research at an in depth level. In the discussion section some papers did not connect or analyze the findings at in depth level, but rather the discussion section was written on a superficial level.

Candidate performance against each criterion

Criterion A: Introduction

The background studies that the students identified were usually well related to their work. Occasionally, students were providing clearly related studies and then simply presenting their hypothesis, without making clear links between the past and present research.

Introductions were consistently developed to show conceptual links to the candidates` own studies. The tendency predominates to analyze the studies without a sufficient attention given to theoretical background. Those few who took both theory and empirical evidence into considerations were awarded by highest mark.

Hypotheses were generally stated clearly, although those with more complex designs tended to be less clearly phrased. Formulation of hypotheses often took the form of a complex sentence with lack of clearly operationalized variables. Another problem occurred with formulation of null hypothesis; negative formulations of experimental hypothesis should be avoided.

Criterion B: Methods: Design

The design was usually well handled by the candidates; however, a number of candidates were not justifying the design, but were defining what is meant by an independent samples design. Justification of the design should refer to the selection being based on a strategic advantage of one of the strengths or characteristics of the design chosen.

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

Criterion C: Methods: Participants

A number of students failed to identify and/or justify their sampling procedure. Additionally, a number of research studies were not specifying their target population or its relevant characteristics. It is important to mention the allocation technique used. It is important to strive for random allocation – especially if an opportunity sample is used.

Criterion D: Methods: Procedure

Procedures were sufficiently described to allow replications. All materials should also be presented in the appendices to allow for 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, noting that the standard deviations seem quite different will allow for some great topics in the discussion section. A surprising number of studies included no measures of dispersion, although most handled the descriptive statistics adequately.

A significant number of studies employed inferential tests that were not appropriate for the data gathered. Complex designs contributed to this problem, as students were unable to deal with the data generated by the design.

It is highly recommended that candidates use one of the inferential statistical tests mentioned in the course guide. These are varied enough to cover the complexity of design expected in this project. They also provide a coherent framework for developing an understanding of the concept of significance quite clearly and easily.

Criterion F: Discussion

The standard 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 previous research studies. They were also able to identify the strengths and limitations of their research and develop logical suggestions for future studies. Unfortunately, a number of candidates struggled with the section. Often failing to identify previous researchers, or just briefly mention just one relevant researcher. All candidates seemed to identify limitations of their research but struggled to identify any of the strengths.

Discussions need to document candidates' competence in both theoretical and methodological issues. Highest scores were awarded for those who showed deep understanding of how theory and methodology are interconnected in psychological research. However, it is quite largely observed that candidates skip relevant theory and take methodological problems as those most important for being discussed.

Criterion G: Presentation

References need to be written in a consistent and standard manner, this includes alphabetical order. The use of secondary sources is appropriate and should be referenced appropriate as such.

It is important to stay within the word limits, as succinct writing skills are important.

Recommendations for the teaching of future candidates

In general, the candidates have presented well-researched and thoughtful experiments. However, some of the candidates need to be made aware of the meaning of a `target` population and also recognize that justification is not the same as a definition.

The design of experiments that will both satisfy the IA requirements and generate data that can be analyzed appropriately by students needs more emphasis. The choices and use of inferential tests needs more guidance. Students should be encouraged to reflect on both the strengths and weaknesses of their studies, and critically engage with the theoretical contexts they base their work on. Attention to the requirements of a standardized referencing system is needed.

It is recommended that candidates choose a topic with a clear independent variable that can be manipulated and that they choose a clear dependent variable that can be easily measured. It is recommended that candidates use in depth analysis in the Introduction to fully explain relevant background research as well as theory that is relevant to the topic. In depth analysis of the findings should also be used in the Discussion Section. Ethical guidelines related to the procedures should be fully discussed. A sample copy of an Informed Consent form should be included with the report. It also recommended that candidates fully explain the justification for utilizing the particular design, sampling technique, and inferential statistical test that was chosen for the experiment.

The operationalization of variables seemed to be one of the biggest problems with students who seemed unaware of the need to have clearly identified variables. Focus on sections varied considerably from school to school; although the Discussion section is a common area where students can usually improve. A greater understanding of the different forms of validity and reliability would help students in their evaluation of research.

Use empirical studies presented during the course to develop each segment of the IA. For example, use an experimental study and require students to identify the aim, hypotheses, IV, DV, design, selection of participants, analysis based on hypothetical data, interpretation of analysis and implication.

Encourage students to identify ethical concerns related to experimentation with human subjects.

Do some exercises during the course of the first year that will help with writing the introduction and discussion section.

The methodological component of the course may need to be introduced earlier, so that students understand the meanings of the terminology within the criteria: target population; sampling techniques; justification for using a particular sampling technique; (parametric vs. non parametric assumptions). Practice in using inferential statistics and choosing the most appropriate test for the design may improve the quality of submissions in this area. Developing a checklist from the criteria for students to check that they have included all components is beneficial for the students.

Ethical considerations - expectations must be clear to students:

It would be useful to focus on the strengths and weakness of different sampling methods and designs and encourage candidates to justify their choice of each in their reports. Proper use of references would also be beneficial.

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

It is satisfying to report that in many centres around the world last year's high standards have been kept. Many candidates showed appropriate knowledge and understanding of the experimental method in their projects by producing relevant reports which did not divert from the main topic, hence the criteria was for most of the time satisfied while giving an opportunity for the candidate to really apply his/her knowledge and understanding of the experimental method to a specific topic. While candidates had very little problems with data collection, a number of them had difficulties with the proper analysis of the data. Also, some candidates failed to successfully complete their tasks because the independent variable was not actively manipulated. The proportion of candidates who are attempting to do a questionnaire or survey still tends to be high.

The majority of candidates had well selected topics, mostly from the domain of learning and cognitive psychology. It was also encouraging to note that the majority of candidates are now elaborating how they are applying ethical guidelines to different aspects of their research study (prior to conducting the research, while conducting the study and when reporting the findings).

Some candidates chose unnecessarily complex experiments to replicate or added several IVs, making their work more difficult and less focused, which invariably led to lower marks.

Candidates from a few centres expressed a deliberate effort not to use any original scientific study as their replication, but instead used their own ideas in order to design an "experiment". At this level such an effort should be strongly discouraged, since it seldom leads to successful results.

Candidate performance against each criterion

Overall, candidates had reasonable understanding of the experimental method. However, some problems that have been appearing year in, year out must be addressed:

- Candidates often fail to clearly state the aim of their research, which gives a general feeling that these candidates are becoming confused when writing about vaguely connected topics.
- Also, candidates at times showed poor knowledge and understanding of the original experiment, which resulted with inaccurate comparison when discussing the findings.

Criterion A

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 is stated rather than the aim.

The majority of candidates were well-acquainted with the findings of the replicated studies, but rarely did they 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 examiners noticed that some students were attempting to write a review of literature but failed to identify and explain one study in detail that has actually been replicated.

Criterion B

Most candidates were able to identify their IV & DV but many could not operationalize them. Candidates especially had problems with operationalizing their IV – many candidates identified the materials that they produced or used as their independent variable (e.g. a list of words, questions stated in the questionnaire, juices of different colour). Candidates were up to the task of identifying the DV, yet a frequent problem was that the dependent variable was not expressed in quantified terms.

Candidates from several schools described their design as a laboratory study (often calling it "experimental") rather than discussing the specific experimental design (independent samples, repeated measures or less frequently matched pairs design).Also, in some cases the experimental design is not always justified. In the majority of cases, the design was often identified correctly but rarely justified in an explicit manner.

The majority of candidates gave some reference to ethical guidelines being followed; if this was not presented in the design section it was usually found in some other sections of the report (usually the procedure) or within the appendices e.g. presentation of the consent form letter or debriefing.

Criterion C

Candidates did a nice job of including appropriate target population characteristics and identifying their sampling technique although a few reported conducting a random sample when they actually used opportunity sampling to recruit participants. Although the use of opportunity samples is appropriate, the justification beyond the fact that individuals were available was not given. It seems that only the strongest candidates justified their use of opportunity/convenience sampling.

Criterion D

The written format of the procedures varied in standard from detailed and accurate to poor. Too many candidates are spending time discussing their production of necessary materials and preparation for the experiment, and then rushing through the procedures that they followed with their participants, therefore leading to a lack of replicability.

In some cases the candidates used a public website to conduct the trials for the study. These candidates did not prepare any of the materials necessary for the replication of the study or

plan their experiment in a consistent manner and therefore cannot be awarded any marks for criterion D.

Criterion E

Without a doubt this was the criterion that was least satisfied. Although candidates usually presented many graphs and tables the presentation of data often did not reflect the aim of the study. In many cases examiners reported that a large number of candidates performed graphing of raw data.

Graphs were computer generated and were generally accurate and clear. There were often cases where the x and/or y-axis was not labelled. Sometimes the labels were not clear "condition 1" rather than the specific condition of the IV.

In general examiners tended to agree that the presentation of descriptive statistics lacked clarity and reflected a lack of understanding of the purpose of statistical analysis of raw data. For example, when presenting descriptive statistics many candidates omitted any type of measurement of dispersion while including all three measurements of central tendency.

In other cases where websites were used to conduct trials the program also included the capability to collect and manipulate data relieving the candidate of any responsibility for data manipulation. In these cases no marks can be awarded as the computation and analysis was not the candidates' work.

Criterion F

Discussions were very variable with some candidates managing to develop them fully and others struggling to do so. The strongest group linked their results to the study they were replicating and identified clear strengths & weaknesses in their methodology. These reports also offered appropriate and insightful suggestions for modification that usually addressed the identified weaknesses. The weaker examples offered very superficial discussion of their results and did not explain the weaknesses or strengths that they identified. Sometimes these samples did not offer any strengths and did not make suggestions for modification.

Weaknesses were more often correctly identified than the strengths of the work. The conclusion was not always presented or not complete and relevant to the study.

Criterion G

Overall, the presentation of the reports was very good. Most candidates included information in the appropriate sections and their abstracts were well written. In general, formatting of references was inconsistent and many reports failed to achieve maximum marks for this. The most common problems related to criterion G were the following:

- The in-text citations were not always correctly given.
- Candidates did not always correctly distinguish between primary and secondary sources.
- Some of the abstracts were not complete.
- Most of the reports seem to be within the set word limits but not all reports gave the word count.

Recommendations for the teaching of future candidates

- There are still some centres in which participants conduct research studies which are not experiments. Teachers should emphasize the importance of actively manipulating an independent variable. In a simple experiment, study participants should be randomly assigned to one of two groups. Generally, one group is the control group and receives no treatment, while the other group is the experimental group and receives the treatment
- Candidates should have the opportunity to develop the skills needed in the production of the Internal Assessment by doing more than this one project. It is recommended to make many types of exercises related to the difficult parts of this project. These exercises could include practice in writing introductions and aims, reinterpretation of findings of published studies, data analysis and analysis of strengths and weaknesses of published studies.
- Care should be taken during the design phase to generate data that lend themselves to descriptive statistics, rather than frequencies or percentages. Candidates should also be reminded to keep the study at a basic level. When original studies are of a complex nature, candidates can conduct a partial replication of the original study with one independent variable.
- The items/questions/video (instrument) used to obtain responses from participants should closely approximate the one used by the researchers whose study is being replicated. These should not be arbitrarily devised by the candidates.
- Since most candidates undertake the research in groups of 3-4, how tasks and responsibilities for conducting the experiment are divided must be explained either in the procedure or the appendix.
- In the participant consent letter and briefing information the actual nature of the experiment should not be divulged, e.g. “this is an experiment in STM recall.” Some specific explanation on how psychologists study memory functions is necessary. Especially in the debriefing more details as to the exact nature of the study should be provided.
- Candidates should avoid labelling the experimental and control groups as “A” and “B” as this may cause some confusion in the analysis and discussion section.
- Candidates should develop skills of analysing the raw data and interpreting findings. Not only does this include how to calculate descriptive statistics but also how to make use of this analysis. For example, candidates should be able to discuss what it means if the calculated means of the two conditions are different, yet the modes are similar, or if the means are similar but the standard deviation of each condition is different. Finally, that marks earned for the Results can be maximized with graphs that reflect the aim rather than raw data.
- Teachers should review the guide to make sure they understand the differences between SL and HL papers. Many students presented a research and null hypothesis, and some carried out inferential statistics. Introductions were sometimes full reviews of literature; some students did not carry out replications, but totally modified studies. Often the cumulative effect was that word counts were exceeded.

- A standardized guide for investigation format and source acknowledgement must be emphasized for next generations of candidates. It is recommended that the reference section should be developed following the standardized format not only using web sources but also books, articles and journals.

Higher and standard level paper one

Component grade boundaries

Higher level

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 4	5 - 8	9 - 12	13 - 17	18 - 23	24 - 28	29 - 52

Standard level

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 4	5 - 8	9 - 12	13 - 17	18 - 23	24 - 28	29 - 44

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

The majority of candidates understood the requirements of most of the questions set for Paper One. Questions set on the cognitive element for Part A have sometimes caused difficulties in previous years, but the question this year received considerable adverse comments from teachers; it was claimed that the gender considerations required were difficult for many candidates to identify. This implied that discussion on how gender considerations could help cognitive psychology understand behaviour was also limited. Apart from this question, concern was also expressed about Part B where two questions were set on applications. While some candidates do find examples of application demanding, others have little trouble in providing a range of relevant uses. Despite these two criticisms the feedback received from most teachers indicated that the paper was fair for candidates and that there was adequate coverage of the syllabus.

Several examiners pointed out that candidates appeared to lack understanding of some of the command terms, particularly the word “compare”. Questions incorporating gender, historical and cultural considerations into their wording tended to provide a general difficulty for certain candidates.

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

SECTION A

Question 1

A successful response to this question requires a clear description of reductionism so that the controversy can be linked to it. While many responses did this, others focused on a description of a specific biological study and failed to address the question. Candidates need to be explicit in identifying the controversy and then in explaining why it is a controversy. Occasionally teachers have queried the ability of candidates to understand what a controversy is, but they should be aware that this word is taken from Learning Outcome 8 in the syllabus for the biological perspective.

Question 2

For this very problematic question there were very few candidates who were able to respond directly to its requirements. While an explanation of gender considerations in relation to the interpretation of behaviour in the cognitive perspective is explicitly required in Learning Outcome 3 of the syllabus, some teachers reported that their candidates had found the question difficult to approach. Several teachers considered the question of gender as too 'peripheral' to the perspective to justify a direct part A question on this topic alone. Despite this caveat some excellent answers were provided by candidates who were able to "think on their feet".

Other attempts at the question often did not address cognitive considerations at all; the frequent accounts of gender differences in behavioural responses to aggressive models in some of Bandura's studies only rarely referred to cognitive factors. Other responses took a speculative view of possible gender differences in research that did not in fact aim to use research gender as a variable: the studies by Loftus on reconstructive memory were often used for this purpose.

More successful answers were generally those that focused on schema theory and studies, including the way in which males and females develop different schemas according to the learning experiences which they obtained during various stages of their development.

Question 3

Better answers presented a clear identification of the method used in the study, an explanation of how the method was used, and a strength and limitation of the method with exemplification from the study. There was a tendency for some candidates to identify the method for the chosen study but then to evaluate the method without reference to the study itself. For example ecological validity was identified as a limitation of a laboratory experiment but rarely applied to the particular study, despite the question injunction for "Using an empirical study..."

Question 4

A number of potentially relevant conditions were offered, but *how* the selected condition contributed to the rise of the humanistic perspective was rarely explained. Better answers explained what was involved in the rise of the perspective, but this did not necessarily mean the first theoretical formulations within it. Many responses discussed a condition that was operative after such beginnings, and really related to a later rise in the popularity or dissemination of the ideas in the perspective, but often this was not acknowledged. The frequently mentioned influence of the Vietnam War is an example of such an influence.

SECTION B

Question 5

Where this question was attempted it tended to produce substantial responses. The focus of the question was intended to include real world applications where the use of research would be for the betterment of psychological behaviour. Candidates were able to choose, for example, drug treatment in dealing with dysfunctional behaviour, or using genetic counselling. Other examples, such as aggression, were also accepted in the context of the question provided that they were clearly linked to the biological perspective. Empirical evidence was usually sufficiently used but often without the necessary evaluation.

Question 6

This question was rarely attempted but the few responses that were presented were mainly successful in the identification of a range of relevant conditions. Explanations were provided on how such conditions played their part in the rise of the cognitive perspective. These explanations referred to conditions such as the reaction against the reductionist approach of behaviourism, developments from psychological thinking during World War II or developments that sprung from analogies to developments in computers and their associated software. Candidates were usually well able to show the positive side of such developments, but these were mainly descriptive and tended to lack evaluation.

Question 7

In this popular choice of question similarities and differences of assumptions were usually well explained. There were times when the wording of the question appeared to be either ignored by candidates or that the meaning of words was not known. For example the concept of an "assumption" was not always understood despite its long usage in IB psychology. "Compare and contrast" also appeared difficult since sometimes only differences were considered. In such cases the maximum score available would be limited to 10 marks. Where the humanistic perspective was chosen for comparative purposes it presented a more difficult exercise than either biological or learning perspectives. The difference between comparison and contrast was not always made explicit. There was also a tendency for some responses to focus on individual studies rather than assumptions for their comparison and contrast; these particular responses did not receive high marks.

Question 8

This was a popular question for HL candidates. There were a few excellent responses and these made strong links between the applications and the humanistic theory with many

answers showing substantial evaluative comment. Although a large proportion of the answers described the applications quite well, including examples such as the nature of open education, motivation in the business world or processes of client-centred therapy, they often failed to support these accounts with relevant humanistic research. Apart from the very good answers, evaluation of any kind was notably lacking.

Recommendations and guidance for the teaching of future candidates

The requirement for evaluation to be included in all answers to Part B questions needs to be emphasised even further. Evaluation was omitted from several essays and examiners were unable to award marks for this critical aspect of candidate writing. Evaluation can take many forms, including critical comment about theories or studies, comparisons with other theories or approaches, the choice of appropriate or inappropriate methodologies, ethical issues, consideration of unclear or non-supportive evidence, and cultural and gender factors that may impact on the ways that behaviour is explained.

Although there is a pronounced tendency for evaluation to be couched in negative terms, it should be remembered that positive criticism can also be an integral part of the evaluative process. It would be a rare research study that requires no evaluation and even where evaluation is mainly negative there may also be elements of the study that become extremely valuable for future generations.

Higher and standard level paper two

Component grade boundaries

Higher level

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 3	4 - 7	8 - 12	13 - 18	19 - 23	24 - 29	30 - 40

Standard level

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

The areas of the program which proved difficult for candidates

Most candidates appeared to be familiar with the format of this paper, but a significant number of students answered more than the required number of questions. Such responses were very superficial. It does not advantage the candidate to answer more than the required number of questions. Candidates are also not recommended to respond to questions from

options that they have not studied – regardless of how interesting questions from other options may look.

Many students had difficulties understanding the concept of application of psychological research, findings, theories, etc. Application is meant to include, for example, practical uses of a theory in a real world setting. For example, research into crowd behaviour in panic situations has lead to changes in the design of aircraft to ensure safety.

At times many candidates seemed to be unable to make their answers relevant to the specific question under consideration. Although very detailed knowledge and understanding may be demonstrated if it is not relevant to the question it cannot be awarded marks. Teachers should endeavour to teach candidates how to tailor their responses to directly answer the question as it has been asked. Many candidates did not seem to appreciate the demands of the questions and just wrote general essays on the topic.

- There was also a compulsion like response to mention gender, cross-cultural and ethical issues without developing anything of substance.
- Misinterpretation of questions seemed to be a universal problem for candidates within this exam paper. Candidates seemed to not "unwrap" the question in order to use time answering what is required, rather than telling what they have learned within their classes.
- Dysfunctional psychology: questions involving methodological issues and comparison between two models of dysfunctional behavior were difficult for them.
- Social psychology seemed to be over generalized. Students were often unable to differentiate between the areas of social psychology covered by the three questions. As in the past, many students had difficulties with understanding what collective behavior is. Inappropriate psychological studies were included in response to this question. Also, many candidates did not have a clear idea of the concepts of conformity, obedience.
- The Psychodynamic question that asked to compare and contrast two models or theories was answered very weakly. Students were very weak at compare and contrast questions. They did not balance the comparisons and contrasting points or weaker students just described the theories per se and did not explicitly put in compare and contrasting points.
- Many candidates struggled simply to answer the question and showed good knowledge related to the question without showing the skills to direct their answers appropriately. Many candidates with weaker knowledge but an ability to direct their answers appropriately did better.

In general, evaluation and analysis was very weak and many answers were just descriptive.

The levels of knowledge, understanding and skill demonstrated

Most candidates anwered questions from Dysfunctional, Psychodynamic, and Social Psychology options. Health and Lifespan were slightly less popular with very few candidates answering questions from Comparative or Cultural Psychology.

The expected full range in quality of responses was seen across this year's sample. There were many outstanding responses read with clear evidence of psychological understanding and clear communication skills. Evaluative skills in the top essays were excellent and were

close tied to the requirements of the question. Evaluation is one skill that has seen improvement in recent years, but still has some room to improve. Now many candidates are memorizing evaluative comments or studies/theories that appear to counter or challenge others. However they may not fully explain these comments or make explicit their influence or link back to what the question is asking. Not all evaluation is relevant to the question.

Some comments from examiners.

- It was pleasing to see that many students seem to understand what methodological and ethical considerations are, as compared to past examinations. The responses to questions on the psychology of dysfunctional behaviour were often strong.
- Some centres showed a high level of knoweledge on psychological theories and studies.
- Many candidates use properly the citation of psychological research within their responses. Some answers showed some descriptive knowledge accompanied by a limited understanding of application and evaluation.
- There appeared to be an improvement with the use of words to demonstrate that similarities and differences were being considered in compare and contrast questions. For example, candidates were writing, `in contrast`, `a similarity`, `on the other hand`, `while`, suggesting that teachers had taken this on board from last year, consequently being reflected in the candidates' improvement in this skill. Knowledge was strongest with 'dysfunctional behaviour' and weakest for questions requiring a consideration of cultural issues.
- Many of the students showed a great deal of knowledge on the general topic area - but did not actually answer the question - they wrote down everything they could about the topic without reading the actual question. So missed out on marks.

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

- Parts A and B type of questions were also confusing for some candidates (questions 7 and 20 for example) where candidates did not know how much detail to add to the answer in parts a and/or what would be appropriate to add in part a vs. part b. Candidates should use the command terms and mark allocations to determine what and how much to write on parted questions.
- Compare and contrast questions also proved a challenge for some candidates, with numerous candidates describing two theories and then concluding with a limited discussion of both similarities and differences.

Dysfunctional Psychology

- Question 7 was the most popular on the entire exam. Schizophrenia and phobias were popular choices. Many students answered Q7 on 'describe one dysfunctional behaviour', some students answered the question very well but some found part b very difficult and didnt analyse the treament programme just described the treatment.

- Question 8 - Mainly interpreted and answered well. Some candidates had problems comparing and contrasting and instead provided detailed description and evaluation of two theories.
- Question 9 - Some good answers here, however some candidates struggled with explaining how the methodological and ethical issues affected diagnosis. This question was not frequently addressed but when it was it seemed to attract the attention of students who knew a lot about Dysfunctional Psychology.

Health Psychology

- The Health Psychology option proved difficult for candidates, where knowledge was very superficial and almost anecdotal.
- Responses to Questio 10 tended to be common sense answers with a few relevant facts. Overly descriptive answers were provided that failed to address clearly the demands of the question (explaining the relationship between gender considerations/differences and interpretation of stress in health psychology).
- Responses to Question 12 tended to be overly descriptive and full of every day examples. Sometimes candidates unsuccessfully attempted to use material from Dysfunctional Psychology to answer this question.

Lifespan Psychology

- Question 13 - Some good responses clearly analysing how early experiences affect adult life. However, many candidates wrote about relevant theory (e.g. Bowlby`s continuity hypothesis) but then did not go on to say how this impacts later life.
- Question 14 - Some handled this well, explaining the strengths and limitations of a relevant method in the context of a relevant study. Weaker candidates evaluated the study rather than the method. In some cases candidates provided answers which reflected only general knowledge of the experimental method or qualitative methods but with no reference to Lifespan Psychology.

Psychodynamic Psychology

- Surprisingly, candidates chose to discuss Freudian applications instead of neo-Freudian ones in question 16 in the Psychodynamic Option. Candidates also had a difficult time knowing what an appropriate application is. In response to question 16 they would write a descriptive account of a neo-Freudian theory without including or evaluating an application.
- Most candidates could find at least one or two similarities/differences to discuss. Some handled this very well - others had similarities and differences but they were not well expanded or elaborated on. Very often candidates provided long and detailed descriptions and evaluations of Freud's theory with a very limited account of one other theory.

Social Psychology

- Regarding question 19, knowing the difference between collective behaviour and individual behaviour in a group proved to be very challenging for most candidates.

Most candidates focused on individual behaviour in a group (e.g., Asch, Milgram). Many very poor answers because candidates discussed conformity studies, particularly Asch's experiment. Most candidates did not seem to understand the meaning of "collective behaviour".

- Question 20. The social question on a study into conformity was a problem with some students answering the question on obedience and just described the Milgram study. Providing evaluation, analysis and appropriate supporting evidence remain a challenge for most candidates in general. For example, in question 20 b, candidates could name some cultural and methodological issues that affect research into conformity without any supporting evidence (i.e. knowledge of studies was weak).
- Question 21. Many focused more on theories of prejudice. Whilst there was an attempt by some to say how this could inform the reduction of prejudice this was not always done particularly effectively. Many essays simply described ways of reducing prejudice, with little evaluation being made.

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

Candidates need to be exposed to the type of questions asked with a lot of emphasis being put on the command terms such as evaluate, compare and contrast, analyse etc. Their answers tended to be very descriptive rather than addressing the question asked. It is important for candidates to have a good understanding of psychological research and theory.,

This extended comment from an examiner is good advice for all candidates and teachers.

- *I saw more answers than usual that clearly were highly practiced, but which were not appropriate answers to the question that was asked. Teachers should stress the importance of responding to the actual question, since a "canned" answer that is not responsive to the question earns few if any marks.*
- *Evaluative comments are needed and students seem to need help in supplementing descriptive accounts with analytical and evaluative commentary. Students need practice in interpreting what a question asks.*
- *Many responses reflect that candidates have done a good job in memorizing relevant information but they are not confident enough to critically discuss this information. Teachers must ensure that candidates develop skills of analysis and evaluation of psychological research and theory.*
- *Candidates should be helped to fully understand the demands of the question as many candidates seemed to possess the knowledge but did not tailor this sufficiently to the demands of the questions.*

Higher level paper three

Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 2	3 - 5	6 - 7	8 - 10	11 - 14	15 - 17	18 - 30

Areas of the programme which proved difficult for candidates

Many candidates appeared to have learned a prepared answer to questions that had previously used key research terms. They struggled to mould their prepared answers to meet the requirements of questions on the paper for this year. In particular the questions on methodological triangulation and on conversational interviews were not always answered well.

Levels of knowledge understanding and skill demonstrated

There was evidence that most candidates had a reasonable knowledge of terms and concepts used in research methods, but their understanding of how this knowledge could be applied in the context of the questions was not always clear.

For example the use of questionnaires in the context of drug use was often answered as a straightforward question on the use of questionnaires in general; conversational interviews were frequently regarded as precisely the same as unstructured interviews.

Strengths and weaknesses of candidates in the treatment of individual questions

Question 1

A minority of candidates mentioned that they had not studied methodological triangulation in considering qualitative methods, but a high proportion of these answers did manage to interpret the term appropriately. Some difficulties were created by candidates who incorrectly used experimental methods as an example of a qualitative approach. In the stress presumably caused by examinations, certain candidates focused on writing about one particular qualitative research method to the exclusion of all other methods. There were erroneous claims that no matter which method was used the end result would always be the same. Very few answers considered the downside of methodological triangulation, but those that did often incorporated a greater understanding of the advantages of this approach. In doing so these answers were often of a higher quality.

Question 2

Many candidates were able to describe conversational interviews and to discuss their strengths and limitations, although these were not always specific to this type of interview. Others had a propensity to describe structured, semi-structured and unstructured interviews as though putting all of this quite detailed information into an answer would gain extra marks.

There were also answers that focused on the ethical issues that each of these methods might have and yet made no reference to conversational interviews.

A definition of conversational interviews was rarely offered although there was mention of their attributes such as being more informal, increased rapport or the general advantage of the more relaxed atmosphere that were engendered. Dire warnings about researcher bias and the credibility of research findings obtained from such unscientific methods were made. Few candidates realised the advantages that are conferred by the mutual expectation of turn-taking by two people engaged in a conversation, or the tendency for the exchange to require more of a role of mutual equality during the conversation. It is this type of trust and respect that makes conversation a rich source of data, provided that it is handled in an appropriate manner by both parties concerned.

Question 3

A substantial number of candidates produced knowledge of the use of questionnaires that was related to the context of the question. They included strengths such as anonymity, the value of a large sample and the relatively straightforward analysis that the data would need. The limitations offered included the potential bias in the answers provided, the problems of a virtually self selected sample, comprehension of the questions and superficial answers. The higher marked scripts used not only many of these points but also referred to the context in which the research was conducted.

In more limited answers the context was either ignored or not adequately addressed. There was a tendency to mention the context superficially in a general phrase such as "……there are advantages and disadvantages in using a questionnaire to study a small community". In many cases no further reference was made to context. In a similar manner the focus of some candidate writing concentrated on the wording of questions and compared this approach with interviews or observations. Candidates also need to read questions with more care. This particular question referred not to drug use but rather attitudes towards drug use.

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

In preparation for Paper three examinations candidates should be taught that all research methods, without exception, have advantages and disadvantages. They should be able to evaluate different research methods in a variety of contexts and applications. Not only will this experience be useful in answering other IB examination papers in psychology but will also help to provide individuals with life time skills.

There is also a need, mentioned by several examiners, that candidates should be able to differentiate between qualitative and quantitative methods and to recognise that each of these approaches has its own merits and demerits. There is a mistaken idea that somehow qualitative methods are inferior to quantitative methods. The choice of which is better used depends upon the research question and context. There is no logic in the suggestion made by several candidates that if qualitative findings are converted to numbers, as with a Likert scale, that they will suddenly confer a scientific respectability upon the findings and the research will be much enhanced because it can be subjected to statistical tests.

Candidates should practise unpacking previous questions so that they carefully consider what they are required to do. The last sentence related to question 3 above, indicates the type of error made by candidates who read questions without due care. Where questions are correctly understood candidates should practice writing responses that are more like discussion or evaluation than merely description. Dividing larger classes into small groups of two, three four or five, and each group to discuss psychological problems or cases can produce dramatic changes for the better where genuine effort is given to the task.