Biology HL & SL

Contents

Grade boundaries

Higher level overall

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 15	16 - 26	27 - 39	40 - 53	54 - 66	67 - 80	81 - 100

Standard level overall

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 14	15 - 25	26 - 38	39 - 50	51 - 61	62 - 74	75 - 100

Higher level internal assessment

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 3	4 - 6	7 - 10	11 - 13	14 - 16	17 - 19	20 - 24

Standard level internal assessment

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 3	4 - 6	7 - 10	11 - 13	14 - 16	17 - 19	20 - 24

Higher level paper one

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 10	11 - 13	14 - 16	17 - 22	23 - 28	29 - 34	35 - 40

Standard level paper one

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 7	8 - 10	11 - 13	14 - 17	18 - 21	22 - 25	26 - 30

Higher level paper two

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 7	8 - 15	16 - 24	25 - 34	35 - 44	45 - 54	55 - 72

Standard level paper two

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 6	7 - 12	13 - 18	19 - 24	25 - 29	30 - 35	36 - 50

Higher level paper three

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 7	8 - 14	15 - 20	21 - 26	27 - 32	33 - 38	39 - 45

Standard level paper three

Grade:	1	2	3	4	5	6	7
Mark range:	0 - 3	4 - 7	8 - 11	12 - 15	16 - 19	20 - 23	24 - 35

Internal assessment

The range and suitability of the work submitted

Many schools presented a very large range of inventive and original investigations. These were a real pleasure to read. Nevertheless, the moderators reported that there appeared to be a heavy reliance in some schools, on classic investigations, some of which are prescribed in the core of the program, with little or no attempt to modify them. Teachers really must coach their students to avoid this problem. This impacts on Personal Engagement

Overall, most of the work was of a suitable standard.

Consideration of safety and ethics were frequently lacking, particularly in work with microbes. The use of body fluids, blood and saliva, was reported by some moderators. This is not authorised at all.

There were some trivial investigations that were not of the appropriate level for the IB biology course and some with little biological content at all, especially when investigations are more in the realm of psychology than biology. A vast majority of the schools did provide appropriate material.

Once again, very few simulations were presented, though the number of databased investigations seems to be increasing a little. Those involving modelling remain rare. Material has been posted on My IB including some exemplars that concern these approaches. We hope that they may clarify their use and how they are marked.

Teachers who physically annotated the candidate's work before uploading, or used the Microsoft Word comments function to annotate electronically submitted work, were most helpful. Examiners found it less helpful when comments were made at the beginning or the end of the work. It was not immediately obvious what the teacher was referring to.

A few samples were uploaded with pages missing. Some were scanned in black and white so it was impossible to understand colour coding used on graphs.

A frequent problem encountered was teachers who did not annotate or comment on work at all (i.e. an unmarked, "clean" copy of the candidate work was uploaded). This made it difficult to follow the motive behind the teacher marks and where possible to support the teacher.

The samples should be completely anonymous. Moderators were still finding candidate names, teacher names, school names and other forms of identification on the uploaded material.

Overall for 66% of schools, the teacher marks were modified by moderation.

Moderators also noted that there were errors of judgement being repeated by schools who had received feedback in previous sessions. It is most important that the feedback gets to the teachers concerned by the school's IB coordinator and that the teachers act upon it.

Candidate performance against each criterion

The application of the assessment criteria by teachers was generally good, though often overgenerous, sometimes very generous. There are cases where teachers are pointing out significant weaknesses in a criterion and then awarding the highest grade. Therefore, more rigour is necessary when applying the final mark. Teachers were occasionally considered too severe.

Evaluation is still the weakest criterion for many. This criterion is difficult and it does discriminate between the candidates. For many moderators Analysis was also a criterion that needed more attention. Many candidates were happy to leave the processing at the level of calculating means and standard deviations.

Personal engagement (PE)

Some form of personal significance was expressed in most cases. While many were clearly inspired by an observation or an issue, many were contrived (for example, "I have always been interested in…"), or there was no expression of personal significance at all.

The originality of the exploration was mostly acceptable and sometimes exceptional. There were, however, too many cases of classic investigations being used with little or no attempt to modify them.

Personal input is evident in the persistence to collect data but also in the research for the background and when establishing the scientific context of the conclusion, in following through the investigation and in the choice of methods of analysis. Once again, this was clearly evidenced for many candidates. For others, it seemed that after a good start with an interesting research question, they failed to follow through.

Personal input can be reflected at the simplest level by having completed the investigation, but those following classic experiments, with no sign of application, cannot expect to score highly. There must be some indication that there is a commitment to the investigation.

A number of moderators observed that teachers seem to be content with a statement of purpose at the beginning of the report and then award a 2. This criterion should be assessed holistically for the entire report, so teachers need to look further for evidence when judging this criterion. A sub-section devoted to Personal Engagement is not what is required. Furthermore, the students will probably need to be taught how to express their personal engagement.

When marking this criterion, teachers should look out for the following:

- A statement of purpose
- The relationship with the real world
- The originality of the design of the method (choice of materials and methods)
- Evidence of trial runs
- The difficulty of collecting data (evidence of tenacity)
- The quality of the observations made
- The care in the selection of techniques to process the data
- The reflections on the quality of the data

- The type of material referred to in the background or in the discussion of the results
- The depth of understanding of the limitations in the investigation
- The reflections on the improvement and extension of the investigation.

Because of its holistic assessment marking this criterion will overlap with components of other criteria.

Exploration (EX)

The research question lacked sufficient focus to obtain the highest mark band for the majority. Scientific names were not always used and the range of the independent variable was frequently not given. For example, a candidate whose question read, "How will different amounts of sugar have an effect on cell respiration in yeast used in bread making?" should have considered including: the species of yeast, the sugar used (was it sucrose, as was assumed?). The word "amount" could have been made more specific by substituting with "mass", or "volume" or "moles". The range of sucrose concentrations to be used should be indicated. A research question can also include how the measurements will be taken by introducing the dependent variable.

It seems from the teacher comments on the samples, that a lot of them seem to be satisfied with less focus in the research question.

The requirements for the background are that it needs to be focused and contain relevant information that is clearly linked to the research question. For example, what was the origin of the pepsin or the milk used in this investigation? There were many cases of superficial or irrelevant material taken from a standard textbook. The independent variable needs to be justified. For example, why was the distance 0-20m from the river used? The dependent variable needs to be explained. For example, how is measuring a change in pH related to the changes in the rate of photosynthesis? The discussion of controlled variables is needed to demonstrate that the student appreciates the other factors that may impact on the experiment. Uncontrolled variables, for example room temperature, may have a significant impact; they need monitoring. One cannot assume that putting the experiments in the same place is enough and setting the air conditioning in a room is often inadequate. Control experiments need to be considered more frequently.

The methods were either written in prose or recipe style. Both were acceptable. Where the method was not clear it affected both the Exploration and Communication criteria. The weaker submissions tended to be from candidates who investigated a topic in which causal relationships are difficult to confirm and a large number of controls are missing. For example, human physiology studies with limited data sets and poorly controlled variables.

When marking this part of the criterion teachers should look out for the following:

- The protocol for collecting the data
- The range and intervals of the independent variable
- The selection of measuring instruments (where relevant)
- Techniques to ensure adequate control (fair testing)
- The use of control experiments
- The quantity of data collected, given the nature of the system investigated

- The type of data collected
- Provision for qualitative observations

Safety, ethics and environmental impact needed to be addressed in a large number of investigations. It is true that some investigations may not have any issues in these areas but there were plenty that did and yet the candidates showed little or no evidence of concern. It is not sufficient to identify potential areas where safety is an issue, there needs to be an indication of how the issue is avoided. All too often the issue of safety is treated without much thought.

There were some microbiological methods being carried out that were very inappropriate for a school environment. For example: collecting bacteria from hands, culturing them at 37°C and exposing them to varying concentrations of antibiotics. This type of investigation is unacceptable.

The following guidelines should be applied:

- Only non-pathogenic strains of microbes from reliable sources such as university laboratories or commercial companies should be cultured. For example, do not culture from hands or swabs of door handles.
- Do not test for antibiotic resistance. There are enough antibiotic resistant strains circulating in the environment without more being selected for.
- Apply strict rules of hygiene and aseptic techniques.
- Do not culture microbes at 37°C. Incubation should be carried out below around 25°C.
- Always label cultured plates so they can be clearly identified and never open them for inspection.
- Tape the lids on but do not tape all the way round a Petri dish. This encourages anaerobic conditions that are best avoided.
- Never assume that what is growing in the culture is the strain that was inoculated, even if non-pathogenic strains have been used.
- Always sterilise used cultures and dispose of the cultures using local health and safety regulations.

The use of consent forms with human volunteers is not systematic. This is an essential ethical practice.

The use of human body fluids appeared as a problem this session with schools permitting students to handle blood and saliva. This is unacceptable practice.

When assessing safety, ethics and environmental issues, teachers should watch out for the following during the experimental phase. Students will need guidance. In the written reports evidence for the consideration of safety, ethical practice and environmental impact can be found as follows:

- Evidence of a risk assessment
- An appreciation of the safe handling of chemicals or equipment (e.g. the use of protective clothing and eye protection)
- Consideration of basic hygiene
- The application of the IB animal experimentation policy
- A reasonable consumption of materials
- The use of consent forms in human physiology experimentation
- The correct disposal of waste

- Attempts to minimise the impact of the investigation on field sites.

Analysis (A)

The presentation of raw data was generally accurate but qualitative observations were often missing. Qualitative observations are expected to accompany the raw data. Their impact will depend upon the nature of the investigation, for example, fieldwork should always have a site description which could take the form of maps, sketches or photographs with annotations. There were investigations that generated only qualitative data e.g. the determination of starch levels using iodine solution by eye. Teachers need to advise their students that this will impact heavily on the Analysis component. The students need to be guided towards more quantitative methods.

Raw data from data logging may be expressed as a graphical readout. It should be accompanied by the necessary information such as units and degrees of precision (if relevant) in the axis titles. These will also impact on the Communication criterion. A candidate should present a representative sample of the raw data, for example, when large amounts of data have been collected using data logging. A representative graphical readout revealing how data is derived is acceptable. In this way the derived data becomes the raw data.

Processing the data varied. Most candidates managed the basics, for example, means and standard deviations, although there were a few that calculated these statistics for everything and anything. Nevertheless, there were still candidates who tried to apply standard deviation to a sample size that was too small ($n<5$). Error bars do not have to be of the standard deviation. Maximum-minimum range bars can be used and this is possible for samples of less than 5.

There were examples of candidates calculating mean rates by averaging the data for all the trail runs and then calculating the mean from this. This is inexact. The rate for **each** run needs to calculated and then the mean from all the rates.

Candidates are still confusing R^2 with the correlation coefficient r. R^2 is the coefficient of determination. R^2 can be used as an indicator of the goodness of fit of a trend line. It can approximate to the product moment correlation coefficient (r) if the trend line is straight, but it is always a positive value unlike the correlation coefficient, which can be negative.

Several candidates were using significance tests from t-test to ANOVA. Although good, they need to be appropriately applied and there needs to be sufficient explanation for the processing to be followed. The use of programmes, such as Microsoft Excel, which produce a statistic, such as a p-value or a correlation coefficient, are fine but the candidate needs to know what the value actually represents.

Basic measurement uncertainties were presented but not discussed. Candidates are expected to appreciate the limitations of their instruments and, where they may have a choice, to select the appropriate one. In biology, the biggest issue for uncertainties is in the variation in the biological material (expressed as standard deviations, standard error or max-min range). Error bars showing variation were frequently used on graphs but their significance, or even what they represented, was often absent. In some cases, the error bars were incorrectly placed or they had no bearing on what

had been calculated. There were cases of students removing outliers from their data during processing. If this is not justified it will be taken as cherry picking the data. This is not good practice.

The interpretation of the data was sometimes well presented after each set of data. Sometimes it was mixed in with the conclusion. In weak candidates the interpretation was a written repetition of the data in the tables with no attempt to point out the trends or to compare data. The use of statistics may have been satisfactory but they were not always well interpreted. As with calculators, the use of a program like Excel is useful but can lead to accepting values without truly understanding them. Huge mistakes can result from this (for example, confusing the t-statistic with the p-value), leading to an erroneous conclusion. Often the interpretation was handicapped by the limited degree of data processing. Some candidates wanted so desperately to support their initial hypothesis that they ignored the evidence pointing to an inconclusive result.

Evaluation (EV)

This was the weakest criterion for many. It is a difficult skill but some candidates just seemed to hurriedly finish off the report. Schools may need to consider the impact of the deadlines for the internal assessment of each subject, theory of knowledge and extended essays on the candidate's workload.

Conclusions were not always supported by the data and explanations were missing. The candidates did not always refer back to their research question at this point. Some candidates were rather overoptimistic in their conclusions. They ignored their processing or did not want to accept their results because they did not "fit". Clearly the data did not fully support it but they would aim to put a positive spin on it. Sometimes a bold statement that the results "prove" the hypothesis right would be made. Few would evaluate the data in the light of a statistical result like the standard deviation. This was a point that clearly discriminated between the candidates but teachers did not always spot it.

An attempt to explain the results in a scientific context is needed for a full discussion and this was frequently superficial or absent. A number of moderators commented that candidates are correctly interpreting statistical significance tests but they are not referring back to the research question.

The evaluation of methodology is still a challenge to most candidates. The consideration of the strengths was frequently missed. Weaknesses were often restricted to practical details or sloppy manipulation and the level of impact on the conclusion was often not discussed. Sometimes the candidates were distracted by less significant issues. Proposed improvements were sometimes unrealistic and often too vague. Extensions were often missed or illogical, not following on from the investigation. Sometimes they were nothing more than improvements to the current investigation. This was an area where moderators felt that teachers were often marking over generously.

When assessing Evaluation of the investigation, teachers should look for the following:

- A conclusion that is supported by the data.
- A conclusion that refers back to the research question.
- An explanation based upon a scientific context.

- A discussion of the strengths – this might be quite general or implicit or it might refer to specific parts that worked well or data that was consistent.
- Discussion of the reliability or the data.
- Identified weaknesses in the method and materials.
- The evaluation of the relative impact of a weakness on the conclusion.
- Sensible, realistic improvements
- Details on the improvements (e.g. not just the investigation needs to be repeated but how many times)
- Realistic extensions that clearly follow on for the investigation.

Communication (C)

The responses to the communication criterion were generally good. Those who communicated well were candidates who had already scored highly in the other criteria.

The most common problems in the work were:

- The use of whole pages for titles. This is unnecessary.
- Whole pages for a list of contents. This is unnecessary.
- Blank data tables presented at the end of the method section (unnecessary).
- Repetitive tables, when one would do. There is often no need for a raw data table AND a table with processed data.
- Raw data relegated to the appendix when there was no reason for it. This upsets the flow of the report.
- Tables split over two pages or with a title on one page and the table or graph on the next.
- Multiple graphs drawn when they could have been combined, this not only saves space but it also improves comparisons.
- Squashed graphs so the distribution of the data is difficult to judge. This is often due to the candidates not reformatting the format of the font.
- Bibliography, footnotes, endnotes or in-text citation missing.
- References with an incomplete format. Often just the URL is given.
- Inefficient data tables headers. The art of designing data tables needs to be taught. A hand drawn sketch of the table layout should be considered first.
- Scientific nomenclature was not always used and the formats were not always respected.

For graphs that result from data logging that are used to derive a value (e.g. a rate) one example can be presented to explain the processing then the rates derived can be organised in a table and it can then be treated as the raw data.

The format for the citations, when they were presented, was mostly correct. However, candidates often missed the need to include the date of access to online citations.

Format of scientific names was sometimes incorrect (small case letter for species name and it ought to be presented in italics).

Units were occasionally missing, or they did not accompany the data.

The use of non-metric units (e.g. teaspoons, cups) was noted by moderators occasionally.

Measurement uncertainties were sometimes missing.

The numbers of decimal places were sometimes irregular, or they did not correspond to the precision of the data.

In general, the reports were of a suitable length.

There were no automatic penalties for reports that were slightly longer, as long as the reports remained relevant and concise. If they were accompanied by extensive appendices, where the raw data is stored, then this would impact on the mark.

Recommendations and guidance for the teaching of future candidates

- Present the criteria to the candidates early on in the course and use them for the assessment of practical work.
- Explain the expectations of each component of each criterion.
- Read and act upon feedback that has been received from previous sessions.
- Ensure that the candidate work has some original purpose. It should not be the repeat of a classic investigation.
- Teachers should add comments throughout the work (rather than at the beginning or end).
- Apply the criteria more rigorously.
- Counsel the candidates on the feasibility of the investigation, focussing research questions, safety ethics and environmental impact, use of statistical programs and the use of citations.
- Teach the students how to express personal engagement in their investigation.
- Teach candidates how to design tables and draw graphs.
- Consider the global context of the candidate's entire IB workload when scheduling the individual investigation in the scheme of work.
- Teachers should visit My IB to see updated examples of individual investigations that are considered adequate (teacher support material).
- Graphs should not be reduced to such a size that they become uninformative, simply to stay within the page limit.
- Candidates should not add on appendices in addition to a write up of about 12 pages and should not send in excessive quantities of raw data from data loggers (although showing an example of how raw data have been processed will be needed).
- Reams of extra work should not be submitted. Teachers marking the work should annotate it if they judge the processed results to be a true reflection of the raw data from, for example, a data logger.
- Full calculations are not expected to be shown, examples will suffice and a worked example from a calculation carried out on a spreadsheet or a programmable calculator will not be expected. However, screen shots should be considered.
- Teachers should ensure that the work is anonymous. The candidate name, the school name, and the session numbers must all be removed before scanning and uploading.

Higher level paper one

General comments

This was a successful paper overall, with many questions that discriminated effectively between candidates and a wide range of coverage of Core and AHL topics. Five questions had very high discrimination indices of more than 0.60, which is a record for this paper and another eight questions had an index of over 0.5. A high discrimination index indicates that more highly achieving candidates are more likely to have answered correctly. There were as ever some problematic questions. Comments in this report will be confined to these and to questions where candidate performance has implications for future teaching. Problematic questions are always regrettable but are almost impossible to avoid in an exam that aims to be innovative with some questions and with an IB rule that pre-testing is not allowed.

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

Question 3

This was an innovative question, but sadly there were ambiguities in it and the discrimination index was very low. Factor II was clearly wrong, as a lower glucose concentration would not give a higher diffusion rate. Answers B and D could therefore be eliminated. Factor I was certainly correct, as initial concentrations would have affected rates of movement across the membrane, but both answer A and C were still possible. The choice depended on whether Factor III was correct, which depended on whether the membrane was permeable to glucose or not. The question did not state this. If the membrane had been permeable to glucose then it seems likely that equilibrium would only have been reached when the levels of fluid had returned to 1, so answer A was considered to be best. 62% of candidates chose that answer, but the poor discrimination shows that some of the stronger candidates had chosen answer C.

Question 5

Question 5 was also only answered correctly by about a third of candidates, though the discrimination index was quite high. The most popular answer was D, which was incorrect – Flask Z did not decompose when Pasteur did this classic experiment. As a result, he showed that exposure to air does not cause decomposition and instead it is the activity of microorganisms that is the cause. The swan neck prevents microbes from reaching the broth.

Question 8

Question 8 was another question that many candidates found difficult, but the high discrimination index shows stronger candidates tended to answer correctly. The commonest mistake was to think that the rate of reaction increases linearly with increasing substrate concentration, but the gradient of the lines on the graph shows that this is not true. Perhaps candidates did not understand what 'increases linearly' means.

Question 9

Questions 9 were well answered with over 90% of candidates choosing the correct answer in question 9 and nearly 90% in question 10. Inevitably with such high success rates, there was little discrimination between candidates, but we can be pleased that candidate performance was strong here.

Question 13

Question 13 was answered correctly by fewer than half of candidates. The word 'live' was printed in bold to emphasise that the homozygous dominant offspring who die before birth should not be included in the ratio.

Question 16

Question 16 had a rather low discrimination index. Part of the reason is probably that the correct statistical test either had been included in a student's teaching and learning or it had not. There was no chance for stronger candidates to work out the answer using their overall understanding of biology. It was pleasing to see that nearly four fifths of candidates did indeed know the correct test.

Question 18

Question 18 was criticised by some teachers who thought that there was more than one correct answer, based on an idea that insectivorous plants are partly saprotrophic and partly heterotrophic. Sub-topic 4.1 of the program includes this Understanding: *Saprotrophs are heterotrophs that obtain organic nutrients from dead organisms by external digestion*. The examining team were in agreement that because insects that plants such as the Australian pitcher plant catch are alive when caught, this is not saprotrophic nutrition and only correct answer is C (autotroph and heterotroph). The statistics for this question do not show any problems, with two thirds of candidates answering correctly, a very high discrimination index, and very few candidates choosing answer A (producer and saprotroph).

Question 19

Question 19 also had a low discrimination index but in this case fewer than half of candidates selected the correct answer. Clearly some of the stronger candidates chose the wrong answer. Answers C and D were the most popular wrong answers. They both include species V, which is adjacent to IV on the right-hand side of the cladogram, but the node where V split from IV is positioned earlier in the cladogram than the nodes separating IV from III and II. Students should be taught to look at the nodes on a cladogram and not the end positions of species to determine relatedness.

Question 20

Question 20 was criticised for expecting candidates to know details of the classification of the Archaea that are not in the program. The only reference to this topic is an Understanding in sub-topic 5.3: *All organisms are classified into three domains*, with this guidance note: *Archaea, eubacteria and eukaryote should be used for the three domains*. The program does not specify that candidates

should be able to distinguish between the three domains by their cell structure, so the criticism of this question is justified – it was based on knowledge beyond the IB biology program. It was fair to expect candidates to recognise that organism X was in domain Eukaryota, but that still left answers B and D and unless candidates had knowledge beyond the program, the only option was to guess between these two answers. More candidates did choose the correct answer B, than D, perhaps because they knew that bacteria do not have proteins associated with their DNA and guessed that the Archaea do have them.

Question 27

Question 27 was criticised for being too difficult and for the 5' and 3' terminals not being indicated. Only 38% if candidates chose the correct answer and the discrimination index was fairly low, so we have to accept that this was not an ideal question. The rational for it was as follows: answers B and C could be eliminated because they contained the base uracil, which is not present in DNA. Answer A could be eliminated because it was the base sequence of the replicated strand, not the template strand. Detailed knowledge of Sanger sequencing was therefore not required, nor did the directionality of the DNA strands need to be known. Perhaps candidates did not remain calm enough and eliminate wrong answers logically in order to arrive at the inevitably correct answer!

Question 32

Question 32 was also criticised for having more than one possible answer. Answers A and D could be eliminated easily because only water moves up a water potential gradient and facilitated diffusion gets substances across membranes, not up stems. The remaining answers were xylem and phloem transport. Candidates were expected to conclude that movement of a fungicide up from the roots to the leaves would be in the xylem in a growing rose plant. Three quarters of candidates chose this answer.

Question 34

Question 34 showed good understanding of the stimulus to flowering in long-day plants with nearly 80% of candidates choosing the right answer.

Question 39

Question 39 was less well answered with only slightly more than 50% of candidates getting the right answer. Answer B was the most popular wrong answer, suggesting two misunderstandings in candidates minds – that the pituitary gland produces estrogen and that estrogen stimulates ovulation directly.

Standard level paper one

General comments

There were very few comments on the G2 forms. One of the comments was that the amount of questions per topic should be balanced. This is exactly what occurs, as the number of questions are proportional to the time allocation per topic. Another comment is on the need of having graphs and diagrams in colour. This is not possible yet. This was a successful paper. The range of topic coverage was wide. Many questions discriminated between candidates and had high discrimination indices of more than 0.40, with 6 questions above 0.54.

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

Question 1

Most candidates realized that cells differ due to the fact that some genes are expressed in some cells and not others. This question discriminated very well.

Question 2

Most candidate recognized the mitochondrion from the micrograph.

Question 3

Unfortunately, this question did not discriminate well. Many candidates failed to realize that a drawing is not an electron micrograph.

Question 4

This question discriminated badly due to the fact that many good candidates believed that the movement of proteins in the mosaic fluid model was due to the presence of transmembrane proteins. The mistake was probably due to a misunderstanding between movement of proteins in the membrane and transport of proteins across the membrane.

Question 5

Common

Question 6

Many candidates confused the process of mitosis with meiosis, therefore answered A instead of B.

Question 7,8

Common

Question 9

This question discriminated very well. Good candidates recognized the sequence of DNA that would produce this amino acid sequence. Unfortunately, many candidates failed to realize they needed the complementary sequence and chose C instead of B.

Question 10

Common

Question 11

This question required higher thinking skills. Candidates had to interpret a graph with three variables. This showed to be quite challenging for many candidates. Although the majority went for the correct answer C, many candidates answered D. They failed to realize that below 200 W m^{-2} the temperature did not affect the rate of reaction, as the graphs are exactly the same for 20ºC and 30ºC.

Question 12

This question was also on meiosis, showing this is a topic that is not well known by candidates.

Questions 13, 14

Common

Question 15

This question discriminated very well. Good candidates realized that the predicted offspring was 50% black to 50% gray. Many believed the probability was 75% to 25%.

Question 16

Candidates really did not need to know about the work in order to answer the question. The idea is for the candidates to understand the possibility of DNA exchange between bacteria. In this question, only *Streptococcus pneumoniae* were used, therefore answer A was not possible and the only possibility was D.

Question 17 to 19

Common

Question 20

This question turned to be a very bad discriminator. Most candidates failed to realize that hair and mammary glands are characteristic to all mammals. One problem could have been they did not realize what the class was, confusing it with vertebrates. Although whales may be observed with hair not all species are born with hair and those that do have hair generally lose it quickly. The hair that does appear on a whale is fairly light and unobtrusive.

Question 21 to 25

Common

Question 26

Although this was a quite discriminating question, many candidates erroneously opted for answer C instead of D. Candidates failed to realize that antibiotics interrupt the metabolism in some pathogens but not others and that pathogens such as viruses have no metabolic processes to interrupt.

Question 27

Common

Question 28

This question discriminated very well and candidates chose the correct answer D

Question 29

This question discriminated very well and was relatively easy.

Question 30

Although this question discriminated very well, many candidates failed to realize progesterone and estrogen are not pituitary hormones.

Higher level paper two

General comments

Comments on Higher Level Paper 2 were received in 39 G2 forms that had been submitted. Most teachers thought that the paper was appropriate in terms of difficulty with about 15% thinking it too difficult. Many teachers felt that the paper was similar in standard to last year's but there was a diversity of opinion over this. 10% of teachers considered it to be a little easier, 34% a little more difficult and 8% much more difficult. The statistics for this paper show a very even spread of marks over nearly the whole of the mark range, with rather more candidates scoring in the upper half of the mark range than last year. The difference between teachers' perceptions of the difficulty of the paper and candidate performance is probably due one part of the paper that looked particularly difficult being an avoidable Section B extended response question. Most candidates did indeed avoid it by choosing to answer the other two questions in Section B.

All but a few teachers considered the clarity of wording and presentation of the paper to be good or excellent.

Only two questions elicited specific comments from more than one or two teachers – 8(a) which was felt to be testing a topic that does not appear in the Core or AHL and 5a where some teachers thought that the diagram of the kidney was potentially confusing. Some teachers felt that syllabus coverage was not as broad as it could have been, with too many marks on ecology and plants, and not enough on human physiology. Paper 2 will always tend to focus on particular parts of the program, to allow candidates to show the depth of their understanding. Inevitably, not all parts of the program can be targeted in this way. Paper 1 with its 40 multiple choice questions is where the breadth of the program is tested with questions from every topic in the Core and AHL.

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

- Understanding the difference between a controlled experiment and an epidemiological survey.
- The difference between pollination and seed dispersal.
- The causes of continuous variation.
- Polyploidy in *Allium* or other groups of plants.

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

- The position of peptide bonds in a polypeptide
- Active transport.
- Sex-linked inheritance of red-green colour blindness.
- Gas exchange in humans.

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

Section A

Question 1: Data based question on the smoking and lung cancer

(a) Most candidates either stated that there was a positive correlation or described the correlation more explicitly. Answers stating that the variables were directly proportional to each other were not accepted because the data is far more scattered than that. The logarithmic x-axis in any case means that direct proportion would not give data points in a straight line on the graph.

(b) Candidates were expected to link changes in DNA to the formation of tumours in their answers. Many scored both marks, though oncogenes were mentioned less often than expected.

(c) The first challenge here was to read off a value from a graph that did not have gridlines – a transparent ruler is useful for this task. Nearly all candidates gave the expected value or something close enough to it. A larger number of candidates failed with the second challenge – quoting the units correctly. The commonest mistake was to state m^{-1} rather than ml^{-1}. In this case one lower case L turns metres into millilitres, so was needed to get the mark, even if it seems a small omission.

(d)(i) Only the strongest candidates were successful here. Nearly all realised that the concentrations would be higher in urine but the reasons given were usually vague. As the mark scheme shows, a reason based on processes in the kidney was required for the mark.

(ii) This was also only answered satisfactorily by a minority of candidates. Most realised that cotinine was the better measure but many gave a reason based on metabolism. The expected reasons could be seen on the axes of the graph -cotinine is in much higher concentrations and has a wider range of values, so it is likely to be a better measure of smoke inhalation.

(e) This was another question where many candidates went wrong. More than half identified the positive correlation in the first graph correctly, but then gave the mantra that correlation does not prove causation. That is true in epidemiological surveys where variables cannot be controlled. The data here came from a properly controlled scientific experiment where the only variable, we can assume, was the dose of NNK. It is therefore quite wrong to dismiss this as a mere correlation – indeed to say that undermines trust in vast amounts of evidence that scientists have accumulated from carefully and properly controlled experiments. This is a most important point which should be stressed strongly. Another weakness in many answers was a failure to understand what the second graph indicated – that smokers do indeed absorb mutagenic NNK from tobacco smoke.

(f) The challenge here was to pick out the significant differences from a rather complicated table of data. Much published research has data that has not been tidied up or made immediately comprehensible, so coping with untidy results tables is a useful skill. Even in news reports, data is sometimes unhelpfully presented so coping with less than perfect data presentation is a skill that is not only of value in science. In this question candidates were expected to state that NNK increased the percentage of mice with tumours and the average number of tumours per mouse.

(g) This question also required candidates to pick out significant differences from a complex results table. There were many valid comparisons that could be made to help construct an answer. A common fault was to base conclusions on differences that were clearly not significant – the standard deviation values for tumours per mouse show that nicotine did not make any significant difference to the number of tumours formed per mouse.

(h) This question proved to be somewhat problematic. It was assumed that candidates would know that nicotine replacement therapy involves giving nicotine in a form other than smoked tobacco. Some candidates thought that NRT is giving a drug other than nicotine to smokers, so their answers did not fit some of the points on the mark scheme. It was still possible to score two marks though for making other valid points, as the mark scheme shows.

Question 2 Protein structure and hydrolysis

(a)(i) This was probably the best answered question on the paper – nearly all candidates could identify a peptide bond.

(ii) In contrast this was a low scoring question. According to the program, primary structure is the sequence of amino acids in a polypeptide, not just amino acids linked together by peptide bonds.

(b) This was well answered and most candidates could give the substrates and products of two hydrolysis reactions in the small intestine.

Question 3 Tandem repeats and DNA profiling

(a) Most candidates successfully gave an answer related to DNA structure as a similarity between tandem repeats and genes, but fewer could give a valid difference other than to repeat that tandem repeats contain repeats.

(b) There was tendency here simply to write about DNA profiling rather than actually answer the question, which was about the role of tandem repeats in profiling.

Question 4 Boreal forests, plant reproduction and photosynthesis

(a) About two thirds of candidates identified Coniferophyta as the dominant plant phylum in the boreal forest illustrated.

(b) Answers were mostly good here as well, with candidates remembering the soil conditions that lead to peat formation.

(c) The challenge here was perhaps greater for examiners than candidates, as there was a huge range of answers. Any valid biological point was credited, but not general ideas about global warming such as sea level rise.

(d)(i) Fewer than half of candidates gave the apparently obvious answer that wind pollination avoids reliance on potentially unreliable populations of pollinating insects. Confusion between pollination and seed dispersal was common.

(ii) There were many good accounts of animal dispersal here, plus frequent reference to the protection that the fruit offers to seeds inside. There was also further confusion between seed

dispersal and pollination from some candidates. Pollination, fertilisation and seed dispersal only get a brief mention in the program but students are expected to know the differences between these processes well enough for them not to muddle them up.

(e)(i) This was generally well answered. The commonest faults were not to give the axis legends fully enough or to reverse them. A few candidates drew an action or absorption spectrum, even though light intensity was mentioned in the stem of the question.

(ii) This was also mostly well answered with candidates able to state the role of rubisco in enough detail to score both marks.

Question 5 kidney structure, osmoconformers and osmoregulators

(a)(i) There were some concerns from teachers that the 3-D nature of the diagram would confuse candidates. Looking at the details of the diagram closely enough, the pyramids of the medulla can be seen and area X is outside them, so must be cortex. Identifying Y also involved more than simple factual recall. Where does the renal pelvis end and the ureter start? The examining team decided that the label was definitely positioned on the ureter side of this junction so only that answer was accepted. Because of the closeness to the word urethra, correct spelling was expected, whereas in other labelling questions some leniency has often been given.

(b) This was a last and rather challenging question in Section A. To their credit, most candidates knew something about osmoregulators and osmoconformers. A few forgot to mention that it is solute concentration that either is, or is not kept constant, or equal to that of the environment. A mark was available for an example of an animal in each group, but few answers included this.

Section B

Nearly all candidates chose question 7 and only about 10% chose question 8. It tended to be very strong or very weak candidates who decided to answer question 8.

Question 6: Active transport, aerobic respiration and gas exchange in humans

(a) Nearly all candidates could outline the process of active transport and strong candidates scored maximum marks in just a few lines. Answers could either focus on calcium uptake or on active transport generally. The commonest error was to call the proteins responsible for active transport *channels* rather than *pumps*. Channels allow passive diffusion across the membrane, not active transport.

(b) Answers tended to be too long here. Many candidates gave a complete account of aerobic respiration, rather than a focussed answer to the question. One way to approach construction of an answer is to think what happens when oxygen is not available. The best answers were concise but still explained fully how oxygen is used in aerobic respiration to allow continued generation of ATP by chemiosmosis.

(c) Nearly all candidates knew something of the human gas exchange system. The best answers focussed on how movement of oxygen and carbon dioxide is maximised by short distances for diffusion and maintenance of concentration gradients. Some candidates think erroneously that pure

oxygen is breathed in by ventilation and pure carbon dioxide is breathed out, or at least the way they phrased their answer suggested this.

Question 7: Sex-linkage, continuous variation and evolution

(a) Most candidates scored full marks for this question or only lost one mark. Generally, it was known that the allele for red-green colour blindness is recessive and located on the X chromosome. The mark scheme rewarded various other points about the inheritance of this condition, including how females can be carriers. One fault in some answers was to refer to a recessive gene rather than a recessive allele.

(b) Candidates did not all read this question carefully enough and as a result there were many answers which gave too much detail about the causes of variation in general and not enough about the causes of continuous variation. The best answers explained polygenic inheritance and the effects of environment clearly and how together they can give a complete range of skin colour or some other continuously variable trait.

(c) The best answers to this question were often the briefest and there were some exemplary accounts of evolution by natural selection. Weaker candidates tended not to understand that natural selection explains the whole process of evolution, instead citing it as one factor that might be involved. Some candidates gave inappropriate emphasis to mutations saying that they are a cause of evolution and that evolution might be particularly rapid if lots of mutations occurred, for example because of radiation after a leak of radioactive materials. Mutation can generate variation, but the pace of evolution depends far more on the intensity of selection and its directionality.

Question 8: Sustainability of ecosystems, roles of the shoot apex and polyploidy

(a) This question was criticised for being based on Option C and although there are two references to sustainability in food production in that option, the idea of the sustainability of ecosystems is a part of 4.1 of the Core: *Ecosystems have the potential to be sustainable over long periods of time*. The real challenge of this question was to base the answer on biological knowledge and understanding, rather than on vague general knowledge. So that candidates were rewarded for any relevant idea, ten points were included in the mark scheme, despite the maximum mark only being three.

(b) This question was answered very well by many of the better-prepared candidates, who were able to include ideas about mitosis, cell division, growth, differentiation, auxin production, phototropism and the production of leaves and flowers.

(c) This justification for this question was the Nature of Science theme in 10.3 of the AHL and also an item in that part of the program: *Application: Speciation in the genus Allium by polyploidy*. Any example of speciation by polyploidy was accepted, plant or animal. Question 8 was answered mostly by very weak or very strong candidates. The former usually knew little about polyploidy and the latter generally understood how polyploidy occurs in general terms, but convincing accounts of an example of speciation by polyploidy were very rare. The majority of candidates who played safe by avoiding question 8 had probably made a sensible decision!

Recommendations and guidance for the teaching of future candidates

- It seems that many candidates do not understand the strength of evidence provided by controlled experiments and confuse this with the correlations that found in data from epidemiological surveys. This distinction is about the most important thing for students to understand in science.
- Candidates should be taught the difference between correlations and direct proportion in relationships between variables displayed on graphs.
- It is wise to bring a transparent ruler to Biology exams as it can be useful when reading values off from graphs.
- It is important to read the whole of each question and then actually answer it. Candidates sometimes write instead about one term that is included in the question.
- Bullet points can be used to write an answer in list form, but they are less good when a coherent explanation is required.

Standard level paper two

General comments

A total of 4755 candidates took the examination: 1253 in English, 3444 in Spanish, 39 in German and 19 in Japanese, representing an increase overall, compared to November 2017.

Of the 33 G2 comments received, 97% thought that the paper was of an appropriate difficulty. 60% thought that the paper was of a similar standard to November 2017, with an almost equal split of the remainder between those thinking it easier and more difficult. In reality, the overall mean mark was higher than last year, with the English language candidates increasing by 2.5 marks, and Spanish language candidates by 0.3. There were 40% G2 comments from English language schools, and 60% from Spanish language schools.

About 90% and 94% respectively said that the clarity and presentation were very good to excellent

All agreed that the paper was accessible to all. There were very few comments from teachers on the G2; most being positive about the question style. Some asked for colour diagrams. Markers did not report that the candidates seemed to be rushing to answer the last questions.

The paper was of an appropriate standard and was fair, while offering some challenges.

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

The command terms such as "discuss" and "evaluate" are poorly applied by many candidates. "Compare and contrast" are not seen as two different answers. Answers frequently did not have enough content to achieve the marking points.

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

Most candidates made good attempts to complete the paper and answer the essay questions reasonably fully. There were few "no responses". A few candidates answered outside the box and there appeared to be fewer extra pages used than usual. The level of knowledge appeared good in some questions, especially on the circulatory system. The data analysis in Section A was quite well attempted.

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

Section A

Question 1

(a) There were many figures outside the range and some candidates, often high-scoring ones, calculated a double percentage change rather than reading off the graph.

(b) There were mainly differences described and surprisingly infrequently the obvious similarity of smoking declining in both genders.

(c) Few candidates attempted to evaluate here, with most stating the trend, and a few seeing the lack of correlation in males around 1960.

(d) There was an overlap with the answers to 1e, and most candidates gained one mark here.

(e) Many candidates repeated answers to 1d in 1e. Only very good candidates were able to evaluate the data. Many confused years since they stopped smoking with year they smoked. Good candidates realized that continuing to smoke increases the chances of lung cancer and that the earlier smokers quite, the smaller the chances.

(f) Most gained a mark.

(g) Many gained 2 marks but some had missed that the question asked for respiratory diseases and chose other cancers.

Question 2

(a)(i) A significant number of candidates failed to label the diagram at all.

(a)(ii) Candidates had to describe the structure of cellulose. This was poorly answered by many candidates. Those that did know some biochemistry scored at least 2 marks.

(b) Generally well answered.

(c) Short/long term stores were frequently mentioned, as well as the greater energy store of lipids.

Question 3

(a) Many answered filicinophytes or evergreens.

(b) A large number of candidates did not link abiotic conditions to not incomplete decomposition. Most only mentioned waterlogged or acidic soils, therefore scoring only one mark.

(c) Most gained a mark by referring to adaptations, and very rarely were 2 marks awarded. Many candidates failed to read the question and answered about the general effects of global warming, even explaining this term, but not referring to its effect on boreal forests.

(d) The majority of candidates labeled correctly, with a sensible curve.

Question 4

(a) Right and left sides of the heart were often muddled but many other candidates labeled correctly.

(b) Many candidates made a good attempt to answer this fully, although the specific roles of elastin, muscle and narrow lumen should be better understood.

(c) There was some knowledge of the SAN and epinephrine; less so for the action of the medulla.

Section B

Question 5

(a) Many candidates scored 4 marks, but a significant number drew eukaryotic organelles.

(b) This was poorly answered in general, with few candidates linking the genome to the proteome - differentiation was recognized without its cause.

(c) Cell theory was well recalled but descriptions of limitations were often incomplete.

Question 6

a) Most candidates scored 3 marks here, without needing to use symbols or a Punnett grid.

(b) There was a wide variety of answers with some candidates barely attempting the question; however, some sound knowledge was also evident.

(c) Some excellent knowledge was shown, with 3 levels of defence described. Some candidates muddle the names and functions of the different white blood cells.

Recommendations and guidance for the teaching of future candidates

Continued examples and practice of command terms are required. Candidates should ensure they have studied all parts of the question on Section B, rather than focusing on one part that they know well.

Candidates should be strongly advised not to answer outside of the boxes provided, but to use extra sheets. They should also not answer Section B questions on the question page.

Higher level paper three

General comments

A total of 1958 responses (term for completed examination papers) were received in English, 276 in Spanish, 11 in Japanese, and 255 in German, for a total of 2300, which represents about 20% of the May 2018 enrolment. Comments on G2 forms were received from 41 teachers (33 about the English paper, 7 about the Spanish, and 1 about the German). Out of these teachers, 34 felt that the level of difficulty of this paper was appropriate, 0 too easy and 7 too difficult. When comparing the difficulty of the paper to last year's, 3 thought it was easier, 26 of a similar standard, 8 a little more difficult, and 2 much more difficult (2 did not answer). About the clarity of the wording, 7 judged it to be excellent, 17 very good, 10 good, 5 fair, 2 poor and 0 very poor. As for the presentation, 11 though it to be excellent, 14 very good, 12 good, 4 fair, 0 poor, and 0 very poor. All these figures tend to show a very slight decrease in satisfaction compared to last year's paper.

There were almost no reports of candidates failing to answer Section A or attempting more than one option in Section B. All responses were scanned in black and white and the individual answer boxes were marked on screen by examiners. Considering this process, a large number responses were very legible; there were nevertheless many, although legible, with very tiny handwriting, and still a limited number of candidates who should make a special effort, especially some German writers using very thick nibs. The vast majority of candidates also comply with the instructions, but there are still some taking the risk of writing outside the prescribed boxes and having part of their answers not being seen on screen.

The majority of candidates in English chose either option C or option D, a lower number option A, and only a small number option B. In Spanish as well, the most popular option was C, followed by D, fewer A and scarcely any B, surprisingly as in N17 a significant number of centres did B, more than the ones in English, and did quite well on it. For German, it was mainly options C and D, with only a few option A.

The responses from candidates writing in Spanish varied more in quality. Particularly notable was the relatively poorer performance in Section A which indicates a lack of knowledge of the required practical skills in the core of the programme, such as of calculations of rates, unit conversions or identification of microscopic structures. Many candidates left questions in Section A with no answer at all. There were also some difficulties in understanding some questions and in expressing concepts.

The biological knowledge of those writing in German still showed major gaps for higher level candidates, who sometimes left some answers blank. Data based questions continue to pose a major hurdle. Often all questions in one section or the (whole) option were interpreted as all relating to the graphs or the initial stem, or it seems that previous exams were not used as part of their learning. Section A caused many problems as candidates had difficulty interpreting the data correctly. The candidates examined in German had more difficulty in answering the questions, but they were more inclined in interpreting the results of data questions than simply describing them compared to last year. Some schools seemed to have neglected to teach the mandatory laboratory experiments. The

conclusions based on the graphs were often inverted, i.e. the cause was seen as the effect. In addition to the apparent lack of biological knowledge it is clear that the candidates' proficiency in German is not at the required level.

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

Many candidates seemed to have difficulty with the areas relating to the practical programme and applications and skills of the syllabus. There was also a major difficulty with basic mathematical skills, mainly dealing with correct decimal positions while converting SI units into others; the concept of rate seems to be unknown for many. The experiment of covering leaves with paper was not well understood. This translates into a poor performance in section A of this examination, but also into difficulties with more practical questions in section B, such as those relating to bioinformatics, usage of laboratory tests, mesocosms and biotic indexes. This seemed even more obvious for those writing in Spanish or German. As always, there is a problem of not reading questions and organizing answers with sufficient care. Some provide answers to past questions and don't realize that new questions may have a different focus; the question about bile production (18 (b)) is a good example of this, where many have described in detail the recycling of erythrocytes by the liver without actually mentioning bile production. Many have a problem with precise terminology, and there are many errors in knowledge of facts and, more specifically, in understanding how elements are interrelated; a good example would be question 8 in which many were confused with the interconnections of neurons within the retina and the transmission of nerve impulses to the brain.

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

The options seemed to be more uniformly covered than Section A, with candidates demonstrating the normal range of knowledge, from excellent to little or absent, but without any specific, outstanding problems with the individual topics covered in this paper. The great majority of the candidates do attempt to answer all questions (except for what we mentioned about the Spanish responses), and the long (last) question for each option was generally well attempted (see details in the next section). Many candidates, although not all, were able to interpret graphs and images well and are better at short theoretical questions. Topics that were generally well answered involve the relationship between brain mass and brain volume, functions of brain areas, biofilms, fermenters, eutrophication, absorption (except bile formation), blood pressure, hormones and blood pH control.

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

Section A

Question 1

(a) Although a good number of candidates provided a correct calculation or answer, there was a lot of confusion within the answers. Many don't seem to know what a rate is and referred to the average volume on the *y*-axis. An answer as number of breaths per second or per minute was expected, but a difference in volume per unit of time was also accepted.

(b) Most candidates have realized that there was an increase in the ventilation rate, fewer providing further explanations. Many candidates did not understand the difference between air volume changes during ventilation and total lung volume.

(c) Most candidates seemed to have guessed incorrectly that emphysema causes lower lung volume (many explaining that less surface means less volume); candidates with a high total score nevertheless had a correct answer.

(d) The role of the two types of pneumocytes was very well known.

Question 2

(a) A large number provided a correct answer, but there are some candidates reversing xylem and phloem, and others providing irrelevant answers.

(b) Although a certain number could correctly calculate 20µm, the number of errors in this question was flabbergasting, showing the lack of knowledge or practice of a basic skill. Besides those who multiplied the measurement by the magnification instead of dividing and those who provided an irrelevant calculation, some measured inaccurately, others measured in centimeters but ignored the units and calculated as millimeters, and others made errors with decimal positions or exponents while converting into micrometers. In Spanish, many students did not even to attempt to answer this.

(c) A good number could provide correct answers, although the diagram (for those who drew one) looked more like a picture taken from a book than what one could observe on a mounted microscope slide. Some answers related more to a root than a stem – although "stem" was clearly stated. Many interpreted "distribution of vascular tissues" in terms of function rather than structure. There was a relatively noticeable number who provided answers so unclear that one could not understand whether the answer applied to a stem or a root.

Question 3

(a) There were many good answers, but too many provided incorrect ones, not being able to figure out what the black card was used to block the light.

(b) Here also, some good answers, but too many not being able to figure out that only the green areas contained chlorophyll.

(c) Most candidates knew that the presence of starch resulted from photosynthesis, but it was difficult for many to discuss whether it was a proof or not that photosynthesis had occurred. Again, in Spanish, many left this blank.

Section B

Option A: Neurobiology and behaviour

Question 4

(a) A large number mentioned the growth of an axon, but many added elements that occurred later in neural development.

(b) Many mentioned the neuron being removed by neural pruning if it was unused, but it seemed difficult to outline other changes.

(c) Although many good answers were seen, many seemed not to understand the idea that the brain could reorganize itself following a trauma.

Question 5

(a) Most candidates could see the increase in brain volume with body mass and realize that the human was an exception.

(b) Most could predict a larger volume and/or an increase of folding for the human cortex and took care not to confuse brain and cortex.

(c) The function of these brain areas were very well known for most candidates, although this was left blank by many in Spanish.

Question 6

(a) Most candidates could describe the data correctly, but fewer related their observations to learned behaviour.

(b) Most candidates deduced that the patient was right handed and provided a valid reason, but there are some who seemed to have confused left and right from the key.

(c) Most candidates knew what operant conditioning and imprinting were, but many, although they could describe a correct example, had difficulty formulating a coherent definition.

(d) A good number of candidates knew that slow acting neurotransmitters trigger the release of secondary messengers, but most couldn't provide a more detailed outline.

Question 7

Most candidates could outline the purpose of these birds' external features or behaviour, but it was more difficult for many to relate it to reproductive success.

Question 8

The majority of candidates could provide correct functions for the rods and cones, but fewer could explain correctly the interconnections between the different neurons within the retina. Many had a vague idea of contralateral processing of nerve impulses, but most couldn't explain it with sufficient or accurate details.

Option B: Biotechnology and Bioinformatics

A limited number of candidates have chosen this option. It is therefore more difficult to generalize about the candidates' performance.

Question 9

Most candidates had a good knowledge of biofilms and the problems they may cause.

Question 10

Most candidates also seemed to have a good understanding of fermenters.

Question 11

(a) Most candidates knew that marker genes were important, but fewer understood that they had to be expressed to determine that gene transfer was successful.

(b) Biopharming seems to be obscure for many candidates, although they know it involves genetic engineering.

(c) Many candidates only have a vague understanding of the usage of A. tumefaciens in producing genetically modified crops. Most answers lacked details.

(d) Many candidates could correctly identify the correct alignment, but blank answers were also seen.

(e) The majority of candidates seem to distinguish between BLASTn and BLASTp and could outline correct reasons, although their answers were not always clear.

(f) Most candidates had difficulty deducing the use of dashes in the alignment.

Question 12

Candidates generally had knowledge of a couple of laboratory tests but had difficulty explaining how they worked and often could not attribute steps to the appropriate test.

Option C: Ecology and conservation

Question 13

(a) The vast majority of candidates stated the correct answer.

(b) A fair number of candidates suggested valid reasons, but it seems that some did not understand that invasive plants are not desirable in an ecosystem or even confused invasive plants with invasive animal species.

(c) Most candidates mentioned competition with native plants, but many didn't go much further; a certain number didn't seem to understand what "control" meant and, as in the previous question part, that invasive plants are not desirable.

(d) Some answers looked more like an outline than a discussion. Most candidates nevertheless understood that invasive species could compete with endemic species, but many answers lacked details. Some even understood that control should be undertaken to preserve the invasive plants.

Question 14

(a) Many mentioned an exchange of some form of matter or energy, but there are some who also mentioned exchange of organisms, which shows that they didn't understand the usage of a mesocosm.

(b) The majority of candidates had difficulty with this question, most of them ignoring that "mesocosms" was plural and identifying only one mesocosm with soil instead of two; other irrelevant answers were also seen.

(c) Most candidates had figured out what the interactions were, but fewer expressed their answer explicitly, outlining clearly how the effect on bacteria occurred.

(d) Most candidates could suggest a valid advantage, but it was difficult for many to suggest a second one clearly.

Question 15

(a) The vast majority of candidates identified group I correctly.

(b) Only a limited number of candidates could identify either n_i, a_i, or both correctly, and many provided irrelevant answers.

(c) The majority of candidates mentioned that they could be used as indicator species, but fewer could provide further details on how.

Question 16

(a) The majority of candidates mentioned that phosphates availability was limited, but it was more difficult for many to address the limitation to agriculture; many included details relating to eutrophication (which was covered by the next question) or tended to repeat answers that would apply to a past examination, not addressing agricultural issues.

(b) The majority of candidates could explain how eutrophication could result from the excessive use of phosphates, but many answers contained inaccuracies, especially relating to the depletion of oxygen.

Question 17

The majority of candidates knew what keystone species were and that they could cause an imbalance of the ecosystem, but fewer could provide clear explanations on the mechanisms involved and quote named examples.

Option D: Human physiology

Question 18

(a) The vast majority correctly stated "small intestine".

(b) The majority of candidates could provide some explanations, but for many these were limited to the breakdown of hemoglobin, as in questions to past examinations, ignoring cholesterol and bile salts, which were specific to the present paper.

(c) The vast majority outlined a valid function.

(d) The majority stated one or more adaptations, but there are some who failed to complete their answer into an explanation by saying what the adaptation was used for.

(e) A large number of candidates could provide a correct comparison for the micelles and lacteals (which were included in the diagram), but the comparison of transport from gut to blood was incorrect for some. A limited number failed to complete two elements for each line, leaving some boxes blank, thus providing an incomplete comparison.

Question 19

(a) The majority of candidates could provide a correct answer about one element, with a good number about a second one.

(b) The majority of candidates had a knowledge of blood pressure and hypertension and their effects on the circulatory system. Many ignored the command term "state" and provided complex explanations, which were irrelevant for marking this question.

(c) The same applied to the measurement of blood pressure and heart rate, although some answers could have been clearer.

Question 20

(a) Many candidates identified the correct elements for this question, but there is a certain number providing incorrect answers, despite what was provided in the diagram. Some candidates do not seem to know what an exocrine gland and/or a follicle is; some didn't know which of the hormones on the diagram were steroids. Some mentioned elements that were not on the diagram.

(b) A large number of candidates provided two correct answers, but some didn't know.

(c) Many could list two correct hormones, but there are some who listed hormones from the anterior pituitary.

Question 21

Most candidates generally provided good, and many complete, thorough explanations, but there were also answers lacking sufficient details or containing diverse inaccuracies. Some candidates ventured sometimes in long Bohr shift explanations, irrelevant for this question which differed slightly from questions in past papers on similar topics.

Recommendations and guidance for the teaching of future candidates

Many recommendations from past subject reports still apply, and are therefore repeated here. Those apply to examination preparation, syllabus coverage, writing skills and examination techniques.

Preparing for the examination

As usually recommended, the use of past papers and markschemes is a valuable tool to prepare candidates for the examination.

Teachers and candidates should nevertheless be warned that it is counterproductive to rehearse or memorize answers from past papers and/or manuals. The extent of the syllabus has limits and it is to be expected that the same topics will be covered throughout the years, but each examination paper is different in coverage and perspective. Past markschemes are useful for familiarization to the paper format, question style, expected vocabulary, depth and variety of elements to include in answers. It is not because a new question shares a few words with a past question that the expected answer will be the same.

Syllabus coverage:

- It is important that schools include the recommended number of hours for a practical programme and internal assessment, in the field or in the laboratory, and include the seven compulsory practical experiments. It appeared, based on their specific answers to questions relating to section A and question 14, that many candidates/schools have not been covering the practical programme, at least sufficiently.
- The syllabus has to be covered completely, including understandings, applications and skills, with links to TOK and NOS when applicable.
- Practical skills and understandings from the core and the AHL are necessary for Section A in Paper 3 and should therefore not be approached as different compartments. Teaching should aim for a comprehensive knowledge of the subject and application of concepts and principles in a wide variety of contexts.
- It is expected that all candidates comply with the mathematical requirements for the programme, including the ability to calculate ratios and deal with decimals. These could be practiced throughout the practical programme.
- A variety of practical examples, data and graphic presentations should be incorporated to the teaching of various topics.
- Teaching of all understandings should be at objective level 3, when applicable.
- Teachers and candidates are encouraged to use multiple sources of reference.

Reading and writing skills:

- Many candidates could have reached a better performance in this examination paper by reading the questions more carefully and including more details in their answers. Something that seems obvious to the candidate cannot simply be assumed by examiners, unless it is explicitly written; this applies to quoting values from graphs without stating their meaning. A school approach to reading and writing into details, using specific vocabulary, could perhaps improve the situation.

- Candidates should be aware that command terms have sometimes a different meaning than they think and should familiarize themselves with them during their course. Teachers should use them throughout the course for their exercises and internal tests.
- Teach the vocabulary of biology as candidates need to use subject-specific vocabulary in their answers. Teachers may choose to build up a glossary of terms used in the programme.

Examination techniques :

- Candidates have a five-minute period before being able to start writing when they are handed Paper 3. They should be aware of this and use this period to carefully read the questions and start mentally planning their answers.
- Many answers could contain more elements and more details. Developing the habit of taking a little time to lay down and organize an answer's core elements would improve answers and prevent omitting important ideas. Coach candidates on how to structure answers: they should take time to consider what is relevant to the answer, leave out what is irrelevant and avoid repeating the same ideas. Encourage candidates to highlight or underline the keywords in the question and plan their answers accordingly.
- It is unnecessary to repeat the question or stem in the answer box; this uses up time and space needed to answer. As can be seen from available markschemes, marking is based on facts and accuracy rather than on style.
- The number of marks indicated in the right margin of the question paper is often an indication of the expected details and number elements for a complete answer. Repeating the same thing many times within an answer doesn't pay much though, elements have to be different. Any "outline" question should never be answered by one word. "Discuss" or "evaluate" questions, including data-based questions, usually require different perspectives to be taken into account.
- With positive marking, candidates are given credit for what they have achieved and for what they have got correct in an answer, rather than being penalized for what they have got wrong, providing there is no contradiction within the answer. Leaving answers blank is therefore not a good strategy.
- Bring a ruler and a square to the exam. This could help measure values on graphs with the required precision.
- Although this did not apply for this paper, respecting proportions in diagrams makes a difference. In all cases, all drawings should be well annotated and labelled carefully.
- Use of colour should be avoided as responses are scanned in black and white to be displayed on line for examiners to mark them.
- Most candidates make a sensible use of continuation answer booklets. The best answers usually fit in the space provided and very few gain additional marks from answers which extend into a continuation booklet. An indication that an answer is continued should nevertheless be made in the main booklet whenever a continuation booklet is used to make sure that the examiner will view it.

Standard level paper three

General comments

Most candidates had a limited understanding of the experiment described in question 1 and were not successful in answering the questions relating to photosynthesis.

Going beyond answering objective one questions was difficult for some candidates, particularly with discuss, explain, suggest or predict commands.

Candidates are often too general in their answers. They must be more specific and use the proper biology terms.

Standard practice of including units was not always followed.

Some "suggest" questions that required going beyond the guide are difficult for candidates. These questions are not something that the students should necessarily have studied in class but together with their knowledge of biology enough information is given for them to arrive at a valid answer. Markschemes for suggest questions are usually very broad as no one answer is expected.

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

The final 4-mark questions at the end of each option were quite well done, as the questions were quite easy and the markscheme broad.

Most candidates scored well on straightforward questions that required no discussion or explanation.

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

Section A

Question 1

Few candidates could outline a reason for inhibiting photosynthesis at the start of the experiment. Many suggested “to maintain leaf colour” or “prove chlorophyll had gone”.

In (b) few could identify the areas showing light and chlorophyll were required for photosynthesis. Many chose X and Y as the area showing chlorophyll as a requirement. Most candidates said that detection of starch was proof of photosynthesis as it is a product of photosynthesis, but this required a discussion looking at other possibilities and very few candidates could do this.

In (c) there were many incorrect graphs of all shapes. Some candidates suggested there would be less photosynthesis or a lower action spectrum in the white areas but very few predicted there would be no photosynthesis in the absence of pigment.

The idea that natural selection would not favour plants with variegated leaves was stated by very few candidates in (d). Some did score the mark for suggesting variegated plants would be less able to compete.

Question 2

The ventilation rate was correctly calculated by the stronger candidates. Credit was given for correct calculations of breaths or volume of gas per unit of time (minutes or seconds).

In (b) most noticed the increase in ventilation rate at the start of exercise, but few explained the change. Most reasonable candidates scored 2 marks here.

The increase in lung volume at rest of a person with emphysema was difficult and probably an unreasonable question. Only a few stated it increased.

Question 3

Most candidates could identify telophase though not many gave a valid reason for 2 marks. Few candidates could calculate the mitotic index correctly.

Option A

Question 4

Many candidates could calculate the metabolic rate of the brain from the data. Most managed one point in (b) for either stating that the muscle uses more energy, or the brain has a higher metabolic rate, but few made both distinctions. Not many candidates gave valid suggestions for why the brain has a high metabolic rate. Most managed to score two marks in (d) for mentioning the medulla and a role such as breathing control.

Question 5

Most candidates could score some marks on these questions. In (a) elongation of the axon was a common answer and in (b) the idea that the neuron would make synapses, migrate or undergo neural pruning was understood. In (c) many correct answers, mostly for brain trauma recovery and some for learning.

Question 6

Many candidates could identify the pituitary gland but fewer the area of the visual cortex. In (b) candidates had a general idea of fMRI but not enough specific detail to explain how it functions. The most common correct answer was for matching performance of task with active area of brain

Question 7

Many candidates labelled the pinna as cartilage or the outer ear. There were some misconceptions about the order of vibrations through the ear; some referred to semi-circular canals and to auditory nerve. The markscheme had enough mark points for most candidates to score well.

Option B

This option was only answered by a few candidates and there were too few responses for any meaningful discussion. The few candidates I saw did not show much understanding in their answers, but a few centres did score well.

Option C

Question 12

The question was poorly understood, and few candidates scored well. Many candidates calculated that the cow would increase by 36 kg in mass given 6 kg of food. Few candidates could discuss sustainability in food production with most simply restating the data.

Question 13

Parts (a) (b) and (c) were an opportunity for many candidates to score marks. Many failed to include the midge larva in (a). Weaker candidates stated keystone species rather than indicator species and gave vague answers for richness and evenness.

Part (d) showed edge effects were not well understood. Some scored a mark for stating that the edge of an ecosystem has different features from the centre. Beyond that there were few correct answers.

Question 14

This was a useful discriminating question which allowed all students to score some marks and the top candidates full marks. Most candidates stated that Japanese stilt grass could be cut in August. Invasive species were fairly well understood but candidates were not always able to understand the questions. Information was given, but not at the right moment. Most mentioned disruption of food chains or competitive exclusion. Many did not realise invasive species are unwanted. In part (d) the most common correct answer was that agents of biological control should not become invasive themselves.

Question 15

Nearly all candidates could state the trend in the graph. The effect of bleaching on corals allowed most students to score marks although some candidates chose to discuss the effect of bleaching on other species in the ecosystem.

Option D

Question 16

Most candidates recognised the loss of skeletal muscle indicated protein was being used aa an energy source though many simply stated the loss of mass from organs Part (c) again showed that candidates have trouble discussing a topic. The guide says that in a discuss question opinions or conclusions should be presented clearly and supported by appropriate evidence. This seldom

occurred. Some candidates correctly stated that the loss of cardiac muscle was indicative of anorexia.

Question 17

Parts (a) and (b) were fairly well answered showing understanding of the defibrillator. Many candidates only scored one point in part (c) for mentioning ventricular contraction.

Question 18

Most candidates could successfully label the epithelial cell layer. Many candidates explained how microvilli increase the surface area for absorption but failed to give further explanation for a second mark. The concept of transport protein was the least answered marking point.

Question 19

In (a) many candidates scored the mark for identifying the Kupffer cell though incorrect spellings were accepted. The most common wrong answer was hepatocyte. Most candidates scored the mark for saying that the iron comes from haemoglobin. Few mentioned phagocytosis and many incorrectly stated that the Kupffer cell transports iron in the blood.

Part (b) discriminated and was well answered by the stronger candidates with most candidates managing to pick up some marks, mainly for identifying both blood vessels and stating that the hepatic artery carries oxygen. Weak answers referred to the hepatic vein.

Recommendations and guidance for the teaching of future candidates

The responses to question 1 on photosynthesis suggest that candidates are not carrying out enough practical work to be able to analyse an experiment and come to a valid conclusion from the results.

Both teachers and candidates need to pay attention to the application and skills listed in the syllabus. Train candidates how to answer the different questions according to the command terms outline, describe, evaluate, suggest, distinguish, discuss etc.

More practice with calculations seems to be needed, including the correct use of standard units.

Candidates should be trained to read the questions carefully and answer the question as it is being asked. They should ask themselves "Do you have to give only examples? Do you have to make comparisons? Do you have to state facts only or do you have to give reasons? Are units required?"

Train students to look at an answer to see if it makes sense. A cow would not gain 36 kg with 6 kg of food so common sense suggest the answer is wrong. Likewise, it is unlikely that we inhale and exhale 6 litres of air per second.

Candidates must select relevant information before they answer the question. They must not write down all they could think of relating to the topic.

Candidates must be to the point. Often, they repeat the question (no marks) and have not enough space for the answer. Also, they often repeat themselves by using other words (no extra marks).

Both teachers and candidates need to pay attention to the application and skills listed in the syllabus. The skill of drawing an action spectrum for photosynthesis is a requirement, so it is surprising that most candidates were not able to perform this task.