May 2014 subject reports

Computer Science

Overall grade boundaries

Higher level

Grade:	1	2	3	4	5	6	7
Mark range:	0–13	14–27	28–36	37–45	46–54	55–63	64–100

Standard level

Grade:	1	2	3	4	5	6	7
Mark range:	0–14	15–27	28–40	41–48	49–58	59–67	68–100

Higher level & Standard level internal assessment

Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0–7	8–14	15–20	21–24	25–29	30–33	34–40

Recommendations for IB procedures, instructions and forms

Generally the work submitted followed the instructions as laid out in the Computer Science Guide and the Handbook of procedures for the Diploma Programme 2014, section B4.4.

Some recommendations from the Principal Moderator:

- one CD should be submitted per candidate in the sample,
- the teacher should check each CD for functionality and content,
- the product folder must contain some evidence of the product – preferably both the final product and the product at design stage, for example both executable jar file and original java files, or an Access database client version that opens a full-screen switchboard and a design

version that does not. In extreme cases, for example when the product is being developed on-line, the product folder should contain screenshots of the product being created.

- the video / screencast should be no more than 7 minutes and should only show the proper working of the final solution – the use of techniques should be described in criterion C using extended writing. It is suggested that students use their Criteria for Success (criterion A) and their test plan (criterion B) to script the screencast.

- In preparation for plans to upload moderation samples online, the completed 4/CompSci form should be added to the CD as a PDF – this typically means 'fill out – print – sign – scan' for each candidate in the sample,

- even though it is not a requirement, teachers are asked to provide pertinent comments on how they awarded marks to the candidates in their sample. This facilitates the moderator's validation of the teacher's marks.

The range and suitability of the work submitted

The described scenarios typically allowed for worthwhile projects. As expected, the majority of solutions concerned programming projects and the majority of those had been coded in Java. On the other hand, an encouraging number of candidates tried their hand at web design, (Access) databases, spreadsheets and Android app design. It is hoped that the range of solutions continues to expand.

The quality of the solutions showed a wide range and not all solutions had been developed to the level of complexity expected of IB exam candidates. Some examples of trivial products include: java programs that mainly focus on GUI and not on actual functionality, Access databases that contain just one or two tables or non-relational tables, websites that are template-based (Wordpress) and that have minimal content. Specifically Access and website solutions tended to be either very strong (incorporating SQL and PHP) or very weak.

Candidate performance against each criterion

A Planning – This was the most straight-forward criterion. However, some candidates did not follow the expected sequence:

- investigate a situation,

- identify client/adviser,

- explicitly consult the client (and/or adviser),

- describe the scenario with reference to the consultation,

- choose a solution,

- describe the rationale for the solution and also for the software to be used,

- outline Criteria for Success for the chosen solution.

Some schools adopted a standard approach where the teacher was the client, setting a task for the student. This is to be discouraged; contrived tasks and clients were routinely seen in the weaker pieces submitted. Too many students had trivial success criteria – these criteria must be specific and testable. The Criteria for Success must also be explicit in the test plan and in the evaluation (and preferably also in the screencast).

B Solution overview – Comparatively this was the worst addressed criterion, and students typically only provided an outline design or even screenshots from the final product (which were discounted). The structured approach of prototyping allowed some students to achieve at a higher level. Records of Tasks were generally only partially complete, either because they did not address all 5 stages (plan, design, develop, test and implement) or because they lacked detail. A wide variety of test plans were seen. The better ones aligned with the Criteria for Success.

C Development – Most candidates made a good attempt to document the development of their product and the techniques used. However, the quality of the explanations and the completeness of techniques typically left something to be desired. The complexity of the product must be justified by the student in the write-up. A seemingly complex product without proper explanations of complex techniques used in the product, only achieves moderate complexity. Similarly, high ingenuity must be justified by algorithmic thinking (e.g. explanations of algorithms or macros).

D Functionality and extensibility of product – The screencast should only show the proper working of the solution as outlined by the Criteria for Success. Many screencasts focused instead on the development of the solution, which made them too lengthy. Others only showed the working of the interface, without showing actual functionality of the intended solution. Some screencasts were in formats that were not recognized by major video software or did not play audio properly.

E Evaluation –For full marks evidence of feedback must be included (typically in the appendix) and it must be referred to in the evaluation against the Criteria for Success. Recommendations should be realistic in relation to the actual product – for example 'adding network capability' is not a realistic improvement for a low-level product.

Recommendations for the teaching of future candidates

The aim of the new IA in Computer Science is to create a working solution for a real client. The consultation (which can be included as an appendix) should be the basis for the description of the scenario, leading to Criteria for Success of a chosen solution.

Criterion B should provide evidence of a rigorous design stage with an overview of all five stages of the project (Record of Tasks), detailed design sketches that include annotations for complex techniques, and a test plan that addresses all Criteria for Success. The best projects included a thorough design stage.

Criterion C provides candidates with the opportunity to show their knowledge and understanding of the tools and techniques used in creating the product. The use of tools/techniques should be explained in relation to screenshots that show their use.

The screencast should be limited to about 5 minutes and should only show the proper working of the final solution. The structure of the screencast should be scripted by the candidate. For example, the screencast could show the testing of the implemented solution following the test plan from criterion B. Successful screencasts showed the solution working with lots of data, but were edited to avoid

viewing tedious data entry. Candidates are advised to test their screencasts on different media players and devices to ensure the playback is correct.

Extensibility is evidenced by a detailed design in criterion B, by a detailed description of the creation process in criterion C and, in case of a programming project, by a properly structured and annotated code listing in an appendix.

Criterion E should provide evidence of a rigorous evaluation stage. The client feedback (in the appendices) should be used to properly evaluate the solution against the Criteria for Success. Recommendations for improvement should go beyond the success criteria that have not been met.

A word of caution: treating the project as a purely academic exercise typically means that there is no proper client and that the solution is not being implemented, which will have an impact on criteria A, D and E.

The recommended word count **for each section**, as indicated in the TSM, is only for guidance. The **overall** word count of 2000 words however, is a fixed limit and a moderator is not required to read beyond this limit, which could cause a loss in marks in criterion E.

Further comments

For additional information regarding the Computer Science IA, please consult:

- Computer Science Guide (pages 56-72).
- Teacher Support Material (Internal Assessment).
- Forms.zip templates.
- Submission of the Computer Science IA in the Handbook for Procedures for the Diploma Programme 2014 (Section B4.4). Note that the Handbook is updated yearly.
- IB Coordinator Notes.

For additional professional development regarding the Computer Science IA, please consider:

- Getting involved in the Computer Science OCC discussion forum.
- Registering for Computer Science workshops (either face-to-face or online).

Higher level paper one

Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0–10	11–21	22–27	28–35	36–42	43–50	51–100

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

Candidates find it hard to focus on discussion questions (for example, Q11, Q12) writing far more that the mark allocation would suggest and without addressing the significant points.

Some candidates did not pay attention to the marks allocated to each question and answer accordingly. Candidates provided many vague, general and ambiguous responses in Section A.

Constructing algorithms proved difficult for many candidates. Responses to questions 14 and 15 ranged from poor to excellent.

The levels of knowledge, understanding and skill demonstrated

Although the syllabus seems to be covered by most schools the performance of candidates hovered around the average.

There are many outstanding candidates and only a few candidates who were very poor in their performance.

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

In general, most candidates seemed to have a good understand of networks and also writing algorithms for one and two dimensional arrays.

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

Question 1

Most candidates were able to identify two features that need to be considered when planning a new computing system.

Question 2

Most candidates correctly explained that beta testing is the last stage of testing performed by end users.

Question 3

Many candidates identified one advantage and one disadvantage of using observations to gather information when planning a new system but failed to explain them.

Question 4

A well answered question. Almost all candidates identified at least one usability issue associated with the design of mobile devices (size of screen, therefore difficult to see; size of keys; battery life, etc.)

Question 5

A well answered question. There were only a few candidates who either confused primary with secondary memory or were not able to identify two types of primary memory.

Question 6

Most candidates provided only a drawing of the resulting linked list. Some candidates failed to explain that the node to be inserted should be initially compared with the head node (node pointed to by head pointer), if not correct position, then move through the list using pointers until correct alphabetical position is found, and then the pointers should be rearranged accordingly.

Question 7

Most candidates knew that a colour represented in a computer would be split into three components. Some did not outline that each component is assigned a certain number of bytes.

Question 8

Many candidates correctly answered this question stating that a peer-to-peer network has no central server and it supports file sharing for collaborative work.

Question 9

Many candidates did not know that paging is used in formation of virtual memory. To increase the amount of primary memory, memory is divided into (tagged) “pages” which are then transferred in and out as required. Some candidates did not attempt this question.

Question 10

Most candidates who answered this question identified autonomy, reactive behaviour, concurrency and persistence as features of autonomous agents.

Question 11

Parts (a), (b) and (d) were well answered by many candidates.

In responses to part c most candidates correctly stated one or two ways in which users can access the functionality of integrated system but often explained only one or provided no explanation/justification at all. Many candidates approach evaluation and discussion questions this way and miss out on marks. For 6 marks, students should make six clear points. The command term “evaluate” requires a consideration of strengths and limitations.

Question 12

Many candidates described well the features of SaaS.

In part (b) most candidates provided vague or too general points in their discussion of the limitation of SaaS in relation to security.

Most candidates, who attempted part (c), stated that extranet is an external extension to a local network with limited access.

In part (d) it was clear that many candidates do not know what a VPN is; this question was not answered well.

Question 13

A well answered question.

Only a few candidates were not able to correctly construct an algorithm to find the highest and lowest frequency of six radio stations and calculate the difference between them.

Question 14

Candidates who attempted this question performed well. Their understanding was good and they could apply their knowledge to this question.

However, it seems that tracing and explaining algorithms is not being taught in some schools. Surprisingly many HL candidates confused branching and looping when constructing pseudocode corresponding to the flowchart given in question 14(a).

Question 15

Questions 15(a), (b) and (d) were where most candidates earned full marks. Sketching the binary tree (part (c)) proved easy for most candidates. Candidates who attempted this question did well. Those who got fewer marks were the ones who did not have enough knowledge about two dimensional arrays.

Recommendations and guidance for the teaching of future candidates

- Tracing and constructing algorithms. Candidates should be exposed to programming concepts and pseudocode. They need to develop their confidence in understanding and writing algorithms.

- Pay attention to the number of marks allocated Usually each correct point is a mark unless specified in the question paper. A question with 1 or 2 marks requires brief answers. Questions with 4,5 or more marks need more in-depth answers suggesting various relevant points.

- Analyze questions in order to determine what is being asked. All responses should contain computer science and not general observations.

Standard level paper one

Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0–8	9–16	17–23	24–29	30–35	36–41	42–70

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

Many candidates struggled with the basic algorithms asked for in questions 12 and 14. Many of the poorer responses suggested a very significant lack of experience putting together pseudocode to describe simple operations such as “find the largest value in an array.”

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

In general, most candidates seemed to have a good understand of networks, how data is transferred across them, and how applications make use of networking.

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

Question 1

This question was generally well-answered.

Question 2

Generally well-answered but many responses were too vague to receive full marks. In particular, many responses failed to identify where in the development process beta testing occurs.

Question 3

Very few candidates earned full marks on this question. While most candidates were able to identify an advantage and a disadvantage, they generally failed to elaborate and explain sufficiently to earn the additional marks. This is an example of where reading the question and considering the marks available can help a candidate structure their response correctly.

Question 4

Some candidates identified an issue that was not related to the design of mobile devices but that arises from the application itself, eg there must be a wireless signal available in order to use the

device. This left them unable to outline a connection to the design of a mobile device and usually resulted in zero marks.

Question 5

Many candidates did not appear to understand the difference between primary and secondary memory and consequently incorrectly listed hard drives, memory sticks, etc as part of their answers.

Question 6

This question elicited a significant number of answers that were incorrect because the candidate confused CASE tools such as WYSIWYG tools for creating application screen designs with CAD tools.

Question 7

This question was generally well-answered.

Question 8

Many of the response to this question were too vague to earn a mark. In particular, a simple statement the computers can act as both clients and servers is not a feature unique to P2P networks. Benefit of the doubt was given where other portions of the candidate's response indicated they recognized that there is no central server in a P2P network, but there were many cases in which the response was simply too vague.

Question 9

Most candidates correctly described the use of data packets, although many responses were rather vague. Relatively few were able to explain that the packet contains information (such as address, byte count, checksum, etc) in addition to the data being transmitted.

Question 10

Most candidates were able to identify several reasons why the speed of a given link might be faster or slower but failed to connect why this might cause the speed across a network to vary. Rather few candidates made the connection with the fact that the packets may travel via different routes and that some routes will be faster than others for the reasons identified. A recurring error made by candidates in many examinations is to confuse speed with duration; a large file will take longer to transmit simply because it is bigger, not because the network is slower when transferring a large file.

Question 11

This question was very poorly answered with most candidates not receiving any marks for their response. The fundamental concepts of data hiding and encapsulation were expressed by only a few candidates.

Question 12

The presentation of this problem apparently confused many candidates as they failed to notice the algorithm they were asked to develop needed to be general and instead chose to "hard code" their

algorithm to the specific truth table given as an example. Even so, many candidates were able to express at least some basic algorithmic ideas in parts (d) and (e).

Question 13:

Responses to parts (a) and (b) of this question were mixed as some candidates were not familiar with the term Software-As-A-Service. Parts (c) and (d) were generally well answered.

Question 14

Part (b) was answered correctly by only a small number of candidates but the remainder or the question was generally well answered.

Recommendations and guidance for the teaching of future candidates

Teachers must teach all computer science candidates to write working programs in a computer language, regardless of which option has been selected. The particular language chosen is of less importance but failure to teach programming as part of the computer science course puts candidates at an extreme disadvantage.

Ensure that candidates consider the context in which a question is asked. When discussing network-based, client/server applications for example, client is NOT synonymous with “customer”.

Ensure that candidates consider the marks available when answering a question; 4 marks for a question that is asking for two issues should immediately suggest that they need to expand on each of those issues and not simply identify them.

Higher level paper two

Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0–10	11–21	22–26	27–31	32–37	38–42	43–65

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

Students had difficulty in answering multi mark questions or ones using words such as Discuss. Students need to be taught how to answer these types of questions by following a couple of simple rules: when stating an advantage it is important to briefly explain why and where there are 2 marks for one aspect the expectation is that the student will make two separate points, or state something and then explain in a sentence.

As an example, it is recommended that teachers spend some time reviewing, with reference to the markscheme, approaches to answering questions such as 13(b) to be awarded the full 6 marks.

It is also vital that students relate their answers to the specifics of the question.

Option A: it was obvious that many students were not well prepared for this option. This was exposed in the first question being poorly answered with students not knowing the fundamental difference between a database and information system. Students were poorly prepared in respect of ER diagrams, normalization, data warehousing, data modelling and data mining.

Option B: Many students adopted a descriptive approach to answer questions about models and variable requirements. It is advised that students adopt a point form or diagrammatic approach. Many students were not aware of neural networks other than at a superficial level.

Option C: Many students appeared to have not been well prepared, possibly this option was seen as only requiring experiential knowledge of the Web; this is not the case and it is expected that all aspects of this option are taught in some detail. Many students did not seem to be aware of specific concepts such as hubs, authorities, deep web, democratic web, ubiquitous computing, the nature of web connections, folksonomies and embedded computing.

Option D: In general this option did not show areas of knowledge lacking in students. There are still students who are not able to answer basic algorithms at an acceptable detailed level – it is expected that students can handle dot notation and be reasonably specific in defining loops, selection and manipulation of objects. Dynamic structures will continue to be examined. Most students could not identify the characteristics of an Abstract Data Type.

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

Option A: Only a small number of students could be said to be well prepared across all questions – it is important that all aspects of this Database option are covered in detail if students are to score well.

Option B: Students appeared to understand the basic idea of simulation and could describe a model or set of inputs. Question 8 was often well answered with students understanding the difference between human and machine languages.

Option C: While some students understood the specifics of terms used in the questions, it is reiterated that general answers and the need for specific knowledge is important if students are to score well. Students generally showed a good understanding of the basic of client-server interaction. The concept of the deep web was understood by a number of students but often there was a lack of detailed understanding; many students showed an appreciation of the ethical design issues related to designers and developers. This pattern was again the case with evolving nature of the web, questions related to searching and understanding of embedded computing. A number of students understood the way that a mother and daughter could interact with ambient intelligence.

Option D: Many students showed a good understanding of UML diagrams, OOP concepts such as inheritance and encapsulation, and variables were well understood. It was pleasing to see that many students could answer algorithmic questions and give sufficient detail – it is important that loops have correct initial and terminating constructs. It should be noted that Dynamic structures will be examined and that often a restricted set of methods will be provided, students will be expected to use these to answer the question.

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

Option A

Question 1

In part (a) many students could not adequately describe the difference between a database and an information system!

In part (b) centralized shared databases are central to world computing, a number of students were not able to describe advantages, the concept was well understood by the better candidates.

In part (c) the query function was identified by many students – however, it was often the case that students did not interpret the situation as read only.

In part (d) concurrency was not well understood.

In part (e) end-user interaction often extended to discussions of customers accessing and updating their personal information, in this case the end-user was the pharmacist.

Question 2

Part (a) was done well.

In part (b) the concept of physical (disk) and logical structures (query in memory) was often not clearly understood.

In part (c) data modelling prior to final design or development was often not fully understood.

Question 3

ER diagrams are detailed in the course outline and in part (a) were poorly attempted by students who seemed to not read the second sentence. This area of analysis and design needs to be addressed.

Part (b) was done well.

Part (c) was done well.

In part (d) creation of a query was not handled well – hence this needs attention by teachers.

In part (e) the concept of a view of the database being a subset defined by a query was not well understood.

Part (f) was not well explained – a view is not part of the physical design of the database on disk.

In part (g) normalization seems to be widely misunderstood.

Question 4

Part (a) was reasonably well answered by better scoring candidates, but often answered as if the student had not heard of the term.

Part (b) was not well answered.

In part (c) many students had some notion of data mining but not necessarily in respect of a data-warehouse. Cluster analysis was not broadly understood.

In part (d) time dependency was not well understood – often being linked to errors rather than simply being able to classify data by date.

Part (e) was reasonably well done, most students would have scored better if their answer was better structured – see markscheme.

Option B

Question 5

Part (a) was well done; inputs and calculations were normally clear, especially where the student adopted a dot point or diagrammatic approach with some explanation.

In part (b) the addition of the taxes and bonds was answered adequately; however, students often neglected to identify the software type and did not reference the consideration of the trends by using the simulation.

Part (c) was well answered by a range of students.

Question 6

In part (a) many students did not address the full range of variables, often missing the score and answer etc.

In part (b) rendering and the need for this in terms of improved visualization were often not answered with sufficient detail.

Part (c) was reasonably well answered, especially by students that adopted a clear layout for their answers.

Question 7

In part (a) data collection techniques were well addressed.

Part (b) was answered reasonably well by a range of students, but often it was not clear that the data from (a) could be feed into some computer simulations and the impacts considered.

In part (c) there often seemed to be a misunderstanding of the way the word criteria was used – the best answers stated levels of congestion and time that would be considered acceptable, and that the results of the model for various sets of inputs would be compared to these statements.

Part (d) was not well done; students often did not consider how the data could be collected, how easily or otherwise this could be done, and then how these aspects could be added to the simulation.

Option C

Question 9

Part (a) was well answered.

Part (b) was reasonably well answered – students and teachers need to review the structure of appropriate answers; it is not sufficient to list or state a way, some explanation is required.

Part (c) was well addressed.

Part (d) was well answered by a range of students, but many did not relate their answers to purchases.

Question 10

In part (a) some understanding of a web crawler was evident, but many students did not start their answer with an outline of how the process starts.

In part (b) the deep web did not seem to be well understood and hence students could not list a reason a web crawler could not find a page generated in this way.

In part (c) many students knew the terms, but often confused them, many demonstrated no knowledge.

Part (d) was handled reasonably well, but often students concentrated on one aspect only.

Question 11

In part (a) a number of students had little idea of the term “ubiquitous computing” and this concept needs to be addressed directly by teachers.

Part (b) was well done.

In part (c) many students knew about grid computing and provided good answers, however, a number simply guessed or simply stated two features without explanation.

Copyright is the right to express or produce, intellectual property is about ownership of the idea, copyright does not protect the idea only the form – copyright is one aspect of intellectual property along with patents, trademarks, trade secrets etc. In part (d) many made this distinction clear, but many confused the terms.

Part (e) was often answered well, but many confused the term democratic web with political and privacy.

Question 12

Part (a) was quite well attempted.

Part (b) was well answered, but many misunderstood.

Part (c) was generally poorly answered, many students simply were not aware of the terms.

Part (d) was well answered by those students who understood the issues related to text and image based searching.

Question 13

In part (a) many students were simply not aware of the term folksonomy. For those that understood, the answers were good.

In part (b) students who understood ambient intelligence provided good answers, but again many were simply not aware of the term.

Option D

Question 14

Part (a) was well answered.

Part (b) was not as well answered, but many students saw that there was a link provided by a variable.

Part (c) was well answered.

Part (d), in general, was well answered, provided the student produced a detailed answer that addressed two ways and included an explanation.

Part (e) was well answered.

Question 15

Part (a) was well answered.

Part (b) was well answered.

Part (c) was well answered.

Part (d) was well answered.

Part (e) was very well addressed by the vast majority of students.

In part (f) the advantage aspect was well answered; the disadvantage less so.

Question 16

Many students addressed part (a) well.

In part (b) the algorithm was well addressed with many students scoring well.

Part (c) asked for the answer to be given in pseudocode, as the logic of the algorithm was key, not the syntax. As with paper 1, all coding, pseudocode and other logical representations were accepted when correct. It is essentially a standard algorithm noting the different entities in a list, but many students struggled with how to approach the problem.

Question 17

Part (a) was very poorly answered; students did not seem to be aware of what ADTs were – the idea that there is a fundamental structure with specific behaviours.

Part (b) was well answered; many students scored full marks and showed good understanding of using a linked list, identifying the end of a list and handling an array.

Part (c) was well handled by a number of students. It is disappointing that a number of less able students made no attempt.

In part (d) there was a pleasing number of students who scored highly on this question.

Recommendations and guidance for the teaching of future candidates

Examination technique – students need to be taught how to answer technical questions using point form or diagrams that are explained. Advantage/disadvantage type questions should have each point explained.

Option A: it was clear that many students had not been prepared; if students are to attempt this option the theory outlined in the subject guide must be covered. The concepts and vocabulary as well as consideration of the way data is stored and manipulated are mandatory.

Option B: the study guide states what needs to be covered; it is strongly recommended that students gain experience of a range of simulation-type problems to identify the inputs, the relationships and how changes in the inputs can be studied to provide study of trends and implications from the simulation. Spreadsheets and other simulation systems available online provide a range of suitable software types.

Option C: teachers need to ensure that the basic theory, terminology and concepts are covered.

NOTE: neither A, B or C can be attempted without exposure to the specific concepts. Being an intelligent user of the web, does not for example ensure that one understands the Web Science related to functioning of the web. It is important to avoid vague superficial answers where there is a well defined body of knowledge that is expected to be known and understood in order to answer questions.

Option D: teachers appear to prepare students well. In general the majority of students showed a reasonable understanding OOPS concepts and the development of algorithms.

Further comments

Some candidates attempted more than one option. These candidates generally performed poorly and it is strongly recommended that candidates maximise their time during examinations by attempting only one option. It was also evident that some candidates chose to complete an option which had not been studied; instead of the one they had been taught. This was the case for both options A and C, which may appear familiar to students but which require much more than superficial knowledge of how to construct and use a database or website. This resulted in some very poor grades for these two options.

Standard level paper two

Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0–6	7–12	13–18	19–21	22–25	26–28	29–45

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

Option A: In general, there was evidence that some of the candidates who tackled this option had not studied it and assumed it to be at a superficial level.

The basic vocabulary of databasing, such as “concurrency”, “user interaction”, “information system” and “data modelling”, appeared unfamiliar to many candidates which resulted in marks being lost. Very few were able to construct an “entity-relationship table” as required and reducing the given table to “third normal form” was rarely completed.

Option B: Although data collection methods were evidently understood, few candidates were able to apply their knowledge to the given situation. Similarly the knowledge of simulation was rarely applied in detail.

Option C: As for option A, there was evidence that some of the candidates who tackled this option had not studied it and assumed it to be at a superficial level.

Few candidates appeared familiar with the terms “hubs” and “authorities”. Many confused the term "democratic web" with the effect of the web *on* democracy.

Option D: Application of knowledge to questions was the only area that caused difficulty to many candidates; for example, outlining the benefits of inheritance in the given situation. A small number of candidates were unable to construct simple code.

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

Option A: Given the quantity of poor responses, there is no one area that can be pinpointed as being well prepared although some candidates showed knowledge in all areas.

Option B: Creating a model and discussing the implications of using simulations demonstrated that students had been prepared.

Option C: As for option A, the number of weak responses masked the excellent preparation of the better candidates.

Option D: Constructing a UML diagram, identifying accessor methods and instance variables, as well as constructing code were all well completed by many candidates and demonstrated that they had been well prepared.

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

Option A:

Question 1

Without reading carefully the stem, many candidates made the mistake of thinking that data on drugs and customers could be updated by the pharmacist, whereas the database was clearly read-only to pharmacist and customer. Consequently some marks were lost in parts (b), (c) and (e).

In part (a) most candidates could outline the function of a database but few understood the meaning of “information system”.

In part (b), apart from those who referred to updating by the pharmacist or customer, the advantages of a centrally based database were correctly identified.

In part (c) “Query” was recognized by many as the function available to the pharmacist.

In part (d) the term “concurrency” was rarely connected to the scenario and often not understood in relation to databases.

In part (e) very few answers to part (e) demonstrated an understanding of database user interaction.

Question 2

This question was related to the construction of relational databases. Responses were often vague, with little understanding of the basic terms.

In part (a)(i) RDMS was generally well defined.

In part (a)(ii) very few candidates referred to the features of the schema and many chose to repeat the term “blueprint” from the stem.

In fact (b) the physical level of the database design was often confused with the actual database.

In fact (c) the concept and importance of data modelling was generally understood although discussions, which should have included individual features to be modelled as well as advantages of modelling these features, were often weak.

Question 3

This question was the most poorly attempted, in some cases showing very little evidence of having followed the course as written in the subject guide.

In part (a) correct entity-relationship diagrams were rare.

In part (b) clear definitions were given by those who understand databasing.

In part (c) almost all candidates identified “Astrophysics” as the answer to this part.

SQL is not compulsory for writing in exams although students will have had experience with constructing queries in whatever database system they used for the course. In part (d) all correctly constructed queries, diagrammatically or otherwise, were accepted for this question which should have been a simple one but often consisted of a vague description.

Parts (e) and (f) were reasonably answered.

In part (g) a correct reduction to third normal form was rare.

Option B:

Question 4

Part (a) was answered well with inputs and calculations clearly outlined.

In part (b) the inclusion of taxes and bond sales was mostly incorporated correctly, and a spreadsheet identified. However, the simulation of a trend over the next few years was often missed or not understood. In particular, changing the values for simulating future effects was often omitted. This had a knock-on effect for part (c).

In part (c) There were some excellent discussions which included advantages and limitations together with a balanced view. Those who did not take advantage of the experimentation nature of simulations found it difficult to relate the simulation to controlling the debt in the future. Full marks required that financial and political implications were included.

Question 5

In part (a) the numbers and operators were correctly identified and most included the way in which they would change throughout the program. The user input, correct answer and score were not all identified and the way in which they would change was rarely detailed enough.

In part (b) some candidates were not familiar with the term “rendering” in relation to visualization.

In part (c) technical implications were a bit weak but there were many thoughtful social implications suggested.

Question 6

In general, not enough time was spent reading and analysing the stem of this question.

In part (a) candidates clearly understood data collection methods but rarely applied them to a situation where traffic flow over a large scale throughout the day was to be measured.

In part (b) a clear model of the infrastructure and the way in which data flow could be applied was often not appropriately described, which is understandable if the traffic flow is based on answers to questionnaires. The use of a spreadsheet would not be considered adequate to represent the infrastructure in this scenario.

In part (c) “Congestion” and “time to cross the city” were the most popular suggestions as criteria to use in evaluating the four methods.

In part (d) the difficulties of including bicycle lanes and improved public transport into the model were often confined to getting the information into the structure and not always into the traffic flow simulation.

Option C:

Question 7

Part (a) was generally a good start to the paper.

In part (b) many answered simply with “encryption” which was only worth one mark; there were **two** ways asked for and they needed to be described.

Parts (c)(i) and (ii) were generally answered well.

In part (d) most understood “server-side scripting” and could discuss the saving of customer data, the sending of related special offers and the use in purchasing goods. However, much of this could have been seen from the question. Very few gave any technical details or reasons for why these events could take place, which is what a discussion would require.

Question 8

This question was disappointing as there were terms that many candidates were not familiar with, despite the fact that they are featured in the subject guide.

In part (a) there was some confusion between “web crawler” and “search engine”. The question also asked for the steps taken to move through the web, not the data collected.

Parts (b)(i) and (ii) followed on from part (a) and only those who had correctly discussed links in part (a) were able to correctly outline a reason for web crawlers not finding the pages.

In part (c) very few candidates were able to demonstrate that they knew the meaning of the terms “hubs” and “authorities”, which are an essential part of the HITS algorithm used by search engines.

In part (d) responses to the discussion question should have addressed the responsibilities of search engine designers and website developers separately and have referred to actions that each should take as well as those they should not take.

Question 9

In part (a) the term “ubiquitous computing” was known by most and many gave suitable examples. Laptops and controlled devices such as washing machines were not credited.

In part (b) almost all candidates could give a limitation to mobile computing.

In part (c) grid computing was understood although many chose to describe it rather than examine two features that would need to be considered before using this type of processing to analyse a large body of data.

In part (d) intellectual property and copyright were understood by practically all candidates.

The concept of a democratic web, in relation to online retailers is that each has the same right to publish and be accessed on the web. In part (e) the discussion should have taken this as a basis and discussed ways in which this right may be taken away by the powers mentioned in the stem and the effect that would have on retailers – large and small. Unfortunately many candidates, unfamiliar with the term "democratic web" understood this question to be referring to using the web to support or deny democracy.

Option D

The scenario for this question was generally well understood although some candidates did not realize that a specimen was a single particular animal. This affected question 11(f) and parts of question 12.

Question 10

Parts (a) and (b) were almost always answered correctly.

In part (c) a significant number of candidates, surprisingly, could not construct a UML diagram.

In part (d) Benefits to the programming team were often vague and repetitive, referring to the benefits of OOP rather than the specific relationships between Specimen, Species and Genus.

Many different answers were given to part (e)(i) which would have been avoided if candidates had read carefully the preceding stem which brought attention to the existence of two separate toString() methods. Despite this, "polymorphism" or "overriding" were often correctly identified for part (e)(ii).

Question 11

In part (a) definitions of "encapsulation" demonstrated understanding but were often lacking in detail.

In part (b) benefits were similar to those in 10(d) and sometimes almost identical, giving general benefits of OOP. A few candidates confused users of the program with programmers.

Parts (c) and (d) were almost always correct except for those who used the Species class instead of the required Specimen class.

In part (e) the code for the Genus class was almost always correct.

Part (f) (especially identifying a disadvantage) was only successfully completed by candidates with a full understanding of inheritance.

Question 12

Answers to part (a) were varied and pinpointed those who had not fully understood the representation of the Specimen class. However, most candidates were able to gain at least half marks.

In part (b) the code was successfully constructed by many candidates, with few errors. On the other hand, there were a few attempts which demonstrated little or no ability to construct code.

Part (c) asked for the answer to be given in pseudocode, as the logic of the algorithm was key, not the syntax. As with paper 1, all coding, pseudocode and other logical representations were accepted when correct. It proved to be a difficult part of the question for a few and was not always attempted. There were also some excellent answers.

Recommendations and guidance for the teaching of future candidates

Option A: When choosing this option it is essential that students have access to a database management system to familiarize themselves with the functions. Any functions required in the examination, such as queries, can be based on this experience.

Case studies, found in most Computer Science and ICT textbooks, can be used to study the use of databases in larger systems from conception to creation. Mini projects can be set where students are required to create entity-relationship diagrams and reduce tables to second and third normal form.

Quizzes and tests can be used to familiarize students with appropriate database vocabulary.

Option B: Students should have hands-on experience of simulating scenarios, and a spreadsheet is suitable for this purpose, although there are also many free programs available that demonstrate simulations. These programs allow the user to change variable values and measure the effect. It is also important to study the use of modelling and simulation in larger scenarios such as those given in B.1.2 in the study guide.

Option C: Ensure that each topic is covered in depth and that students are familiar with the correct terminology. One way to approach this option is to get students to research specific topics, such as search engines and then to present their findings. There is a lot of interesting information on the Internet and students are likely to be motivated to find out as much as possible. Obviously, a session summarizing and tying up would need to be given by the teacher, especially to ensure that the terms and vocabulary referred to in the subject guide are fully understood.

For options A, B and C, ensure that candidates practise answering examination papers and pay particular attention to following:

- Reading and interpreting the stem of a question so that answers are relevant to a given scenario.
- when the command term “discuss” is used, give relevant points of view as asked for in the question, such as advantages and disadvantages when appropriate, and complete with a synthesis.

Option D: The two parts, theory and coding practice both need to be fully covered and are best approached by integrating them into practical examples. A coding exercise can be shown each time a feature such as encapsulation is covered, followed by a class or homework question which covers the same feature. Definitions and advantages can then be discussed and noted.

Further comments

Some candidates attempted more than one option. These candidates generally performed poorly and it is strongly recommended that candidates maximise their time during examinations by attempting only one option. It was also evident that some candidates chose to complete an option which had not

been studied; instead of the one they had been taught. This was the case for both options A and C, which may appear familiar to students but which require much more than superficial knowledge of how to construct and use a database or website. This resulted in some very poor grades for these two options.

Higher level paper three

Component grade boundaries

Grade:	1	2	3	4	5	6	7
Mark range:	0–4	5–8	9–11	12–14	15–17	18–20	21–30

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

The new Higher level paper 3 is based on a case study that is made available to schools a year in advance of the May session in which it is examined. A new case study is presented each year. The case study differs considerably from its format in the previous specification in that it now demands an in-depth study of a specific area that is both current and relevant to computer science. Failure to carry out this research would severely compromise the student when sitting the paper, as answering the questions with no more than general knowledge will gain few if any marks.

Although this new nature of the case study was stressed in the new subject guide and in the many workshops that took place in the preceding two years, it was clear that many schools had failed to devote sufficient time or had not gone into the topic at a sufficient depth.

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

The two definition questions (Question 1) were reasonably well answered showing that the additional terminology had been looked at.

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

Question 1 will always refer directly to terms listed in the “Additional Terminology”. Understanding of these terms is necessary as they will encounter 2 mark questions, and a simple 1-line answer may not successfully address two marking points. Repeating phrases given in the case study (e.g. “fingerprint of acceptable applications”) is unlikely to gain any credit unless expanded on further.

With the publication of the case study a year in advance there is an opportunity to set preparatory work for the students. If they were to arrive back at the start of the second year with a good understanding of the new terminology, then a solid foundation will already have been made.

Question 2 looked at both IPS systems and DoS attacks. These questions will be more involved, demanding detailed answers. As in all the questions on this paper, marks are awarded for showing an understanding of the issues and for the use of correct terminology. Hence for part (a) marks were awarded for naming a type of malware / describing the data that is sent out / the fact that the malware entered undetected / the role of the IPS in detecting unusual patterns / excessive traffic / blocking or reporting the transmission etc.

The DoS question referred to the three types of attack specified in the case study. It was also emphasized that the person named in the study had to prepare a report on these attacks, yet many students were still unable to explain how any of the three were carried out. Prevention methods mentioned were rather generic in most cases, such as firewall and IPS, which again indicates that in-depth research had not been carried out.

Question 3 will be of increased difficulty requiring quite detailed knowledge of one or two areas referred to in the study. In this paper the related issues were the SSL protocol and MitM attacks. Most students gained marks for aspects of MitM (being part of the conversation / unknown to client or server / manipulation of the data sent etc.). Marks were also awarded for showing knowledge of the SSL protocols and also for any credible means of attack against the protocol (false certificates / intercepting of initial request / brute force on older versions etc.).

The most important question in terms of marks is question 4 (12 marks). This discussion question will bring together different aspects from the study that are directly related to one of more of the challenges detailed towards the end of the study. For this paper, it was the BYOD challenge that was tested. The students need to first identify the different areas that need to be discussed. In this case they were:

- the reasons for deciding to operate the policy
- the dangers of putting such a policy into operation
- countermeasures that could be taken

Credit was only given for answers that referred directly to BYOD. Answers that dealt only with security issues in general tended to end up in the lower mark band (1–3). Those that did deal with BYOD issues but on a main descriptive level would achieve the 4–6 band. The 7–9 band was reached by students who showed understanding of relevant issues and the top band (9–12) for those who shown a detailed understanding of all three issues.

Although there were some excellent responses, overall the answers for this question were disappointing and showed quite clearly that the required level of research had not been carried out.

Recommendations and guidance for the teaching of future candidates

- Set work during student vacations, as the study is published in May each year.
- Research to a depth that is appropriate for a HL IB course.
- Be aware of how the questions are structured (and will remain structured), and how they relate to different sections of the case study.

- Research the "Challenges Faced" with particular care, as the 12 mark question will always relate to one or more of these challenges.

Further comments

It is appreciated that the research required puts significant demands on both teachers and students. However, 30 hours have been allocated for this element, which is same amount of time that is allocated for the selected option (paper 2). It is also anticipated that most of the research will be carried out by the students themselves outside of class, either individually or in groups as directed by the teacher. It is hoped that this element of the HL course will provide the opportunity to study an area of computer science that is both current and interesting.