Biology subject report

2023 cohort

February 2024

Queensland Government | QCAA Queensland Curriculum & Assessment Authority

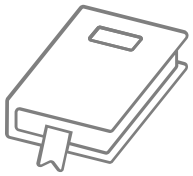
240033

Queensland Curriculum & Assessment Authority
PO Box 307 Spring Hill QLD 4004 Australia

Phone: (07) 3864 0299
Email: office@qcaa.qld.edu.au
Website: www.qcaa.qld.edu.au

Contents

Introduction



Throughout 2023, schools and the Queensland Curriculum and Assessment Authority (QCAA) continued to improve outcomes for students in the Queensland Certificate of Education (QCE) system. These efforts were consolidated by the cumulative experience in teaching, learning and assessment of the current General and General (Extension) senior syllabuses, and school engagement in QCAA endorsement and confirmation processes and external assessment marking. The current evaluation of the QCE system will further enhance understanding of the summative assessment cycle and will inform future QCAA subject reports.

The annual subject reports seek to identify strengths and opportunities for improvement of internal and external assessment processes for all Queensland schools. The 2023 subject report is the culmination of the partnership between schools and the QCAA. It addresses school-based assessment design and judgments, and student responses to external assessment for this subject. In acknowledging effective practices and areas for refinement, it offers schools timely and evidence-based guidance to further develop student learning and assessment experiences for 2024.

The report also includes information about:

- how schools have applied syllabus objectives in the design and marking of internal assessments
- how syllabus objectives have been applied in the marking of external assessments
- patterns of student achievement.

The report promotes continuous improvement by:

- identifying effective practices in the design and marking of valid, accessible and reliable assessments
- recommending where and how to enhance the design and marking of valid, accessible and reliable assessment instruments
- providing examples that demonstrate best practice.

Schools are encouraged to reflect on the effective practices identified for each assessment, consider the recommendations to strengthen assessment design and explore the authentic student work samples provided.

Audience and use

This report should be read by school leaders, subject leaders and teachers to:

- inform teaching and learning and assessment preparation
- assist in assessment design practice
- assist in making assessment decisions
- help prepare students for internal and external assessment.

The report is publicly available to promote transparency and accountability. Students, parents, community members and other education stakeholders can use it to learn about the assessment practices and outcomes for senior subjects.

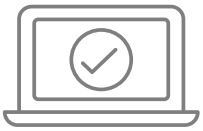
Report preparation

The report includes analyses of data and other information from endorsement, confirmation and external assessment processes. It also includes advice from the chief confirmer, chief endorser and chief marker, developed in consultation with and support from QCAA subject matter experts.

Subject highlights

441
schools offered Biology

82.1%
of students completed 4 units



99.16%
of students received a C or higher

Subject data summary

Subject completion

The following data includes students who completed the General subject or Alternative Sequence (AS).

Note: All data is correct as at January 2024. Where percentages are provided, these are rounded to two decimal places and, therefore, may not add up to 100%.

Number of schools that offered Biology: 441.

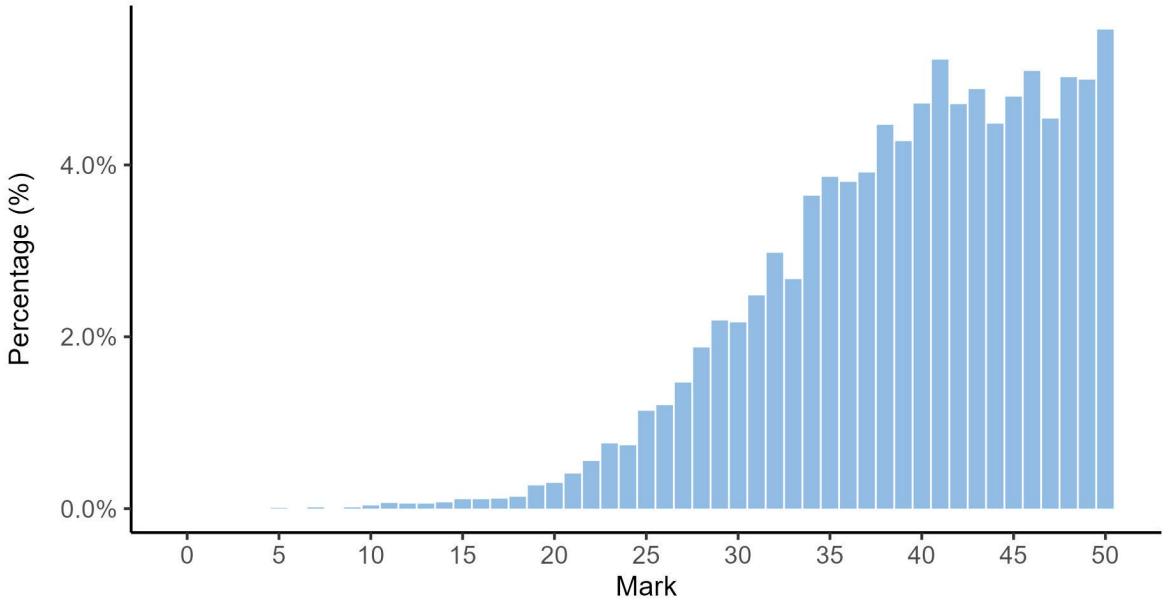
Completion of units	Unit 1	Unit 2	Units 3 and 4
Number of students completed	16,606	15,326	13,634

Units 1 and 2 results

Number of students	Satisfactory	Unsatisfactory
Unit 1	15,438	1,168
Unit 2	13,875	1,451

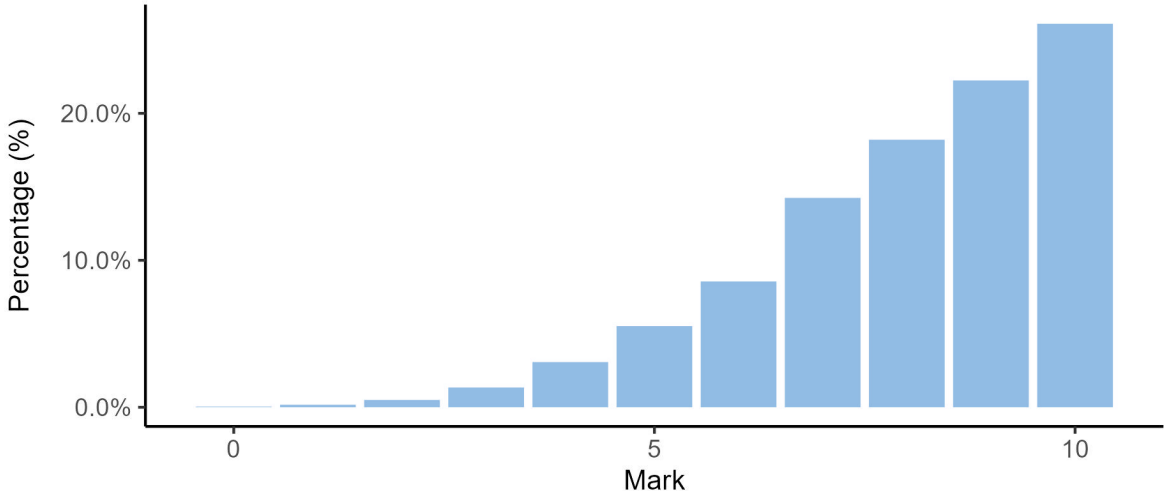
Units 3 and 4 internal assessment (IA) results

Total marks for IA

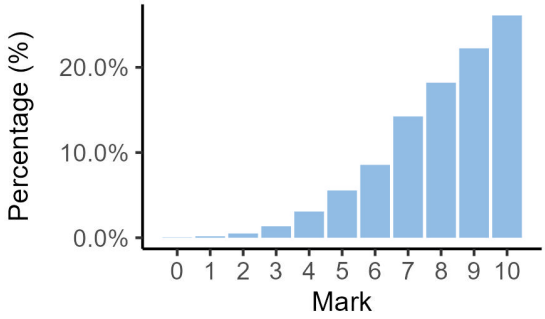


IA1 marks

IA1 total

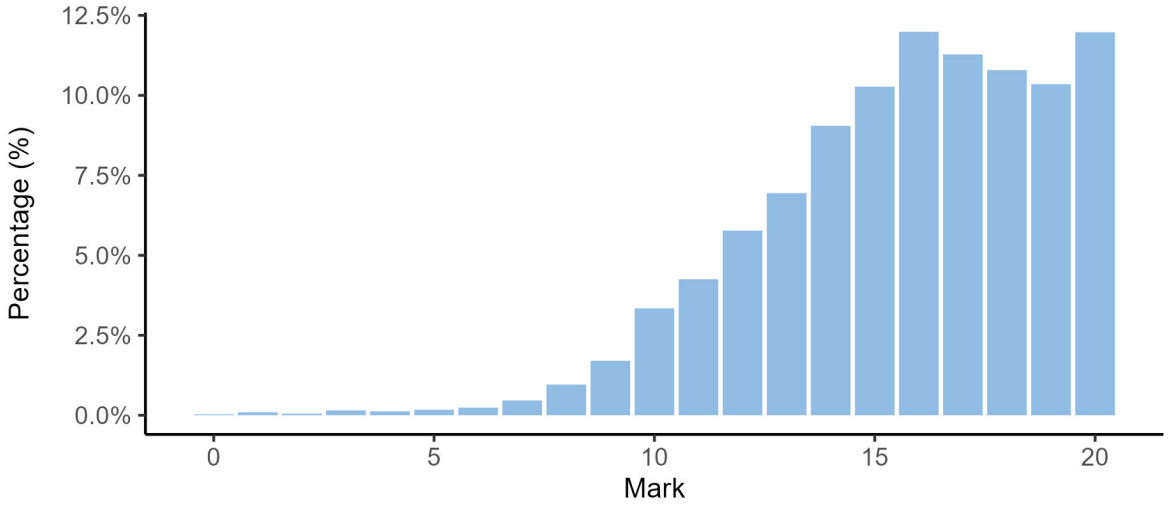


IA1 Criterion: Data test

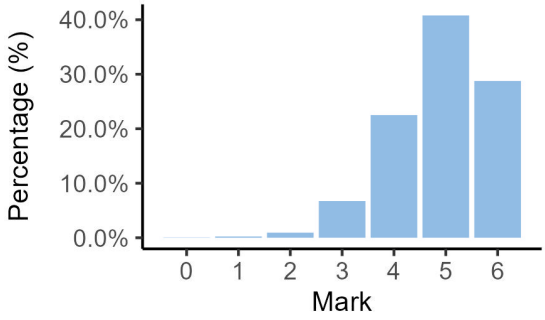


IA2 marks

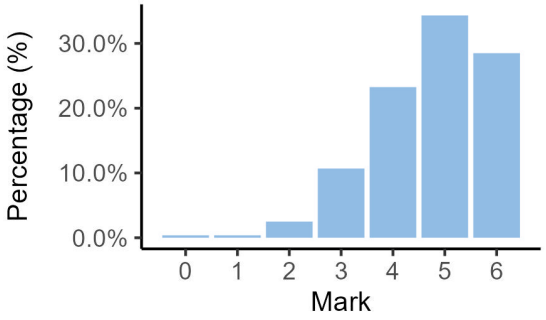
IA2 total



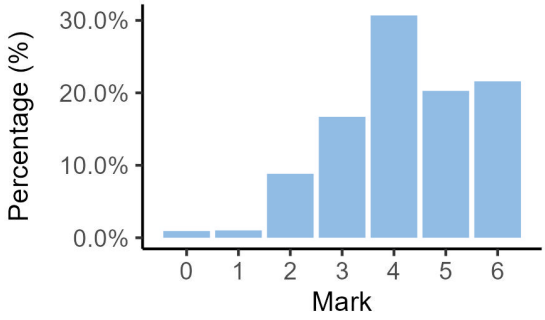
IA2 Criterion: Research and planning



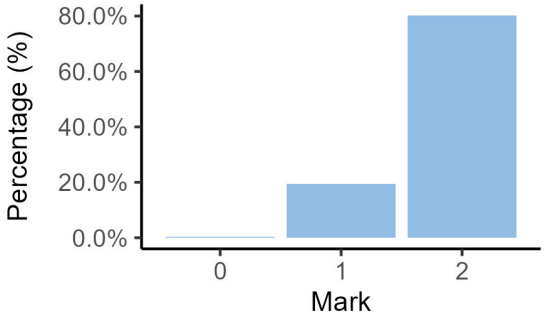
IA2 Criterion: Analysis of evidence



IA2 Criterion: Interpretation and evaluation

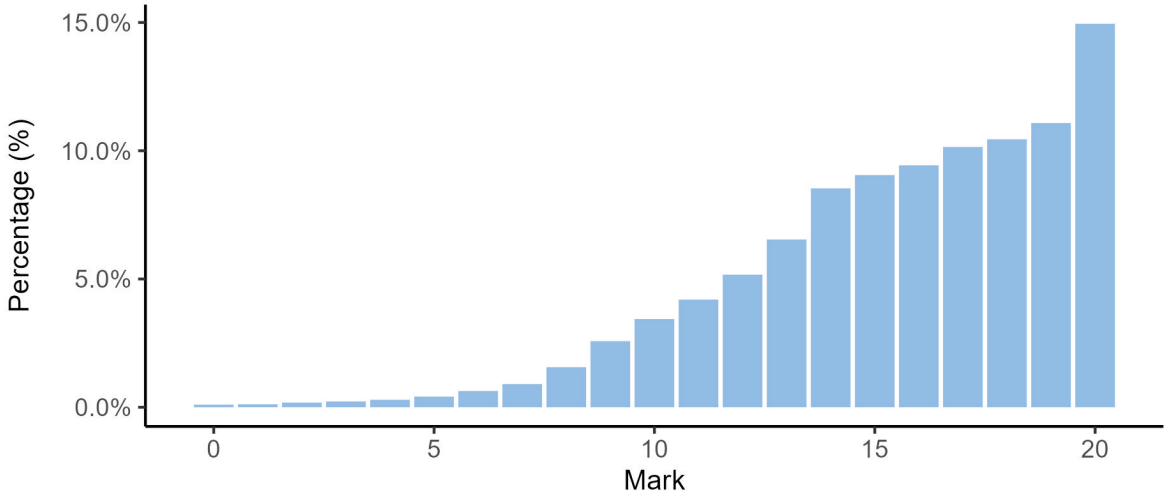


IA2 Criterion: Communication

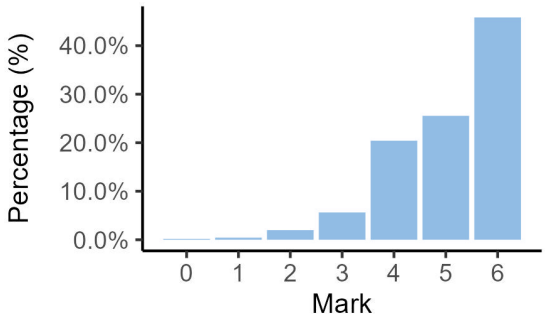


IA3 marks

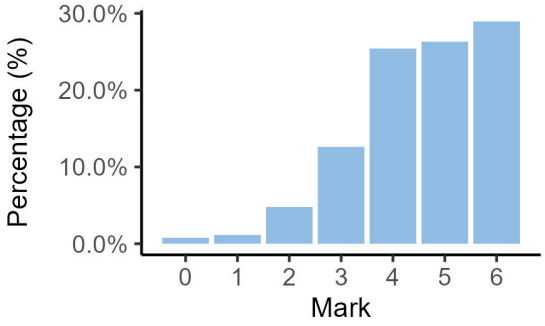
IA3 total



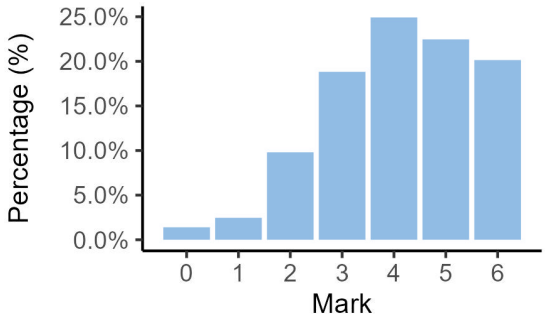
IA3 Criterion: Research and planning



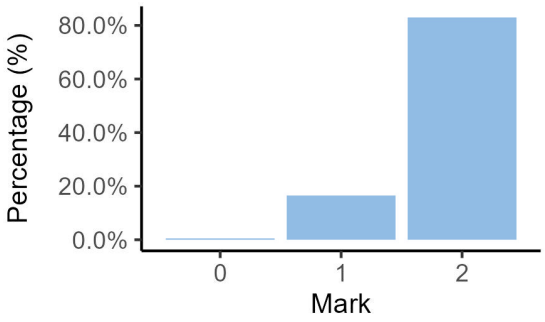
IA3 Criterion: Analysis and interpretation



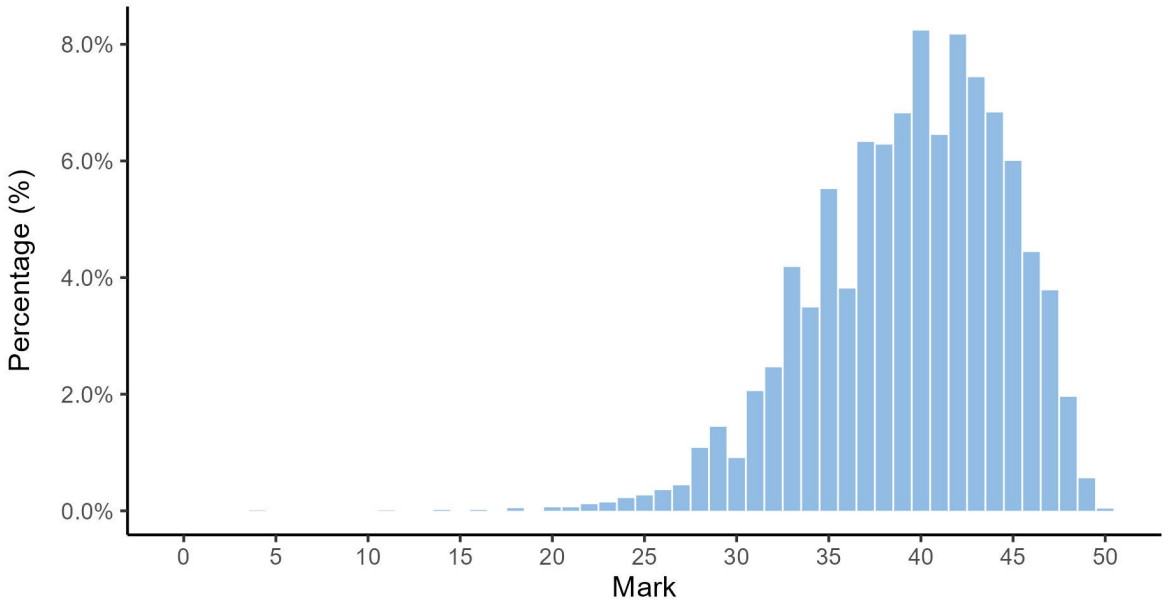
IA3 Criterion: Conclusion and evaluation



IA3 Criterion: Communication

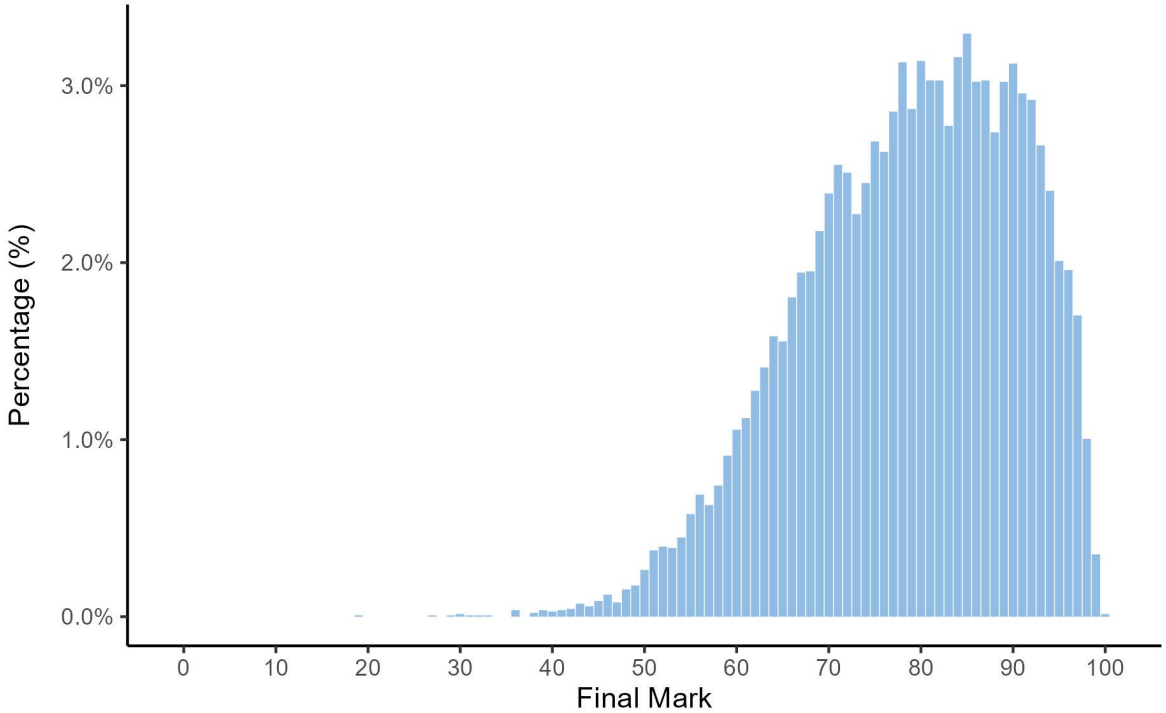


External assessment (EA) marks



Final subject results

Final marks for IA and EA



Grade boundaries

The grade boundaries are determined using a process to compare results on a numeric scale to the reporting standards.

Standard	A	B	C	D	E
Marks achieved	100–86	85–70	69–49	48–19	18–0

Distribution of standards

The number of students who achieved each standard across the state is as follows.

Standard	A	B	C	D	E
Number of students	4,488	6,103	2,928	115	0

Internal assessment

The following information and advice relate to the assessment design and assessment decisions for each IA in Units 3 and 4. These instruments have undergone quality assurance processes informed by the attributes of quality assessment (validity, accessibility and reliability).

Endorsement

Endorsement is the quality assurance process based on the attributes of validity and accessibility. These attributes are categorised further as priorities for assessment, and each priority can be further broken down into assessment practices.

Data presented in the Assessment design section identifies the reasons why IA instruments were not endorsed at Application 1, by the priority for assessments. An IA may have been identified more than once for a priority for assessment, e.g. it may have demonstrated a misalignment to both the subject matter and the assessment objective/s.

Refer to *QCE and QCIA policy and procedures handbook v5.0*, Section 9.6.

Percentage of instruments endorsed in Application 1

Number of instruments submitted	IA1	IA2	IA3
Total number of instruments	439	439	435
Percentage endorsed in Application 1	48%	88%	91%

Confirmation

Confirmation is the quality assurance process based on the attribute of reliability. The QCAA uses provisional criterion marks determined by teachers to identify the samples of student responses that schools are required to submit for confirmation.

Confirmation samples are representative of the school's decisions about the quality of student work in relation to the instrument-specific marking guide (ISMG), and are used to make decisions about the cohort's results.

Refer to *QCE and QCIA policy and procedures handbook v5.0*, Section 9.7.

The following table includes the percentage agreement between the provisional marks and confirmed marks by assessment instrument. The Assessment decisions section of this report for each assessment instrument identifies the agreement trends between provisional and confirmed marks by criterion.

Number of samples reviewed and percentage agreement

IA	Number of schools	Number of samples requested	Number of additional samples requested	Percentage agreement with provisional marks
1	434	2,660	0	95.85%
2	433	3,343	320	84.99%
3	433	3,320	165	84.53%

Internal assessment 1 (IA1)

Data test (10%)

This assessment focuses on the application of a range of cognitions to multiple provided items.

Student responses must be completed individually, under supervised conditions, and in a set timeframe.

Assessment design

Validity

Validity in assessment design considers the extent to which an assessment item accurately measures what it is intended to measure and that the evidence of student learning collected from an assessment can be legitimately used for the purpose specified in the syllabus.

Reasons for non-endorsement by priority of assessment

Validity priority	Number of times priority was identified in decisions*
Alignment	135
Authentication	0
Authenticity	0
Item construction	49
Scope and scale	70

*Each priority might contain up to four assessment practices.

Total number of submissions: 439.

Effective practices

Validity priorities were effectively demonstrated in assessment instruments that:

- provided 2–4 datasets that were
  - related to Unit 3 (or AS Unit 1) subject matter
  - based on practicals, activities or case studies the students had experienced in class, e.g. mandatory and suggested practicals
  - of appropriate scope and scale, i.e. not overly complicated, free of distractors, and used minimal explanatory text
- contained a series of short response questions that
  - required students to use the datasets
  - allowed students to demonstrate a range of cognitions aligned to objectives 2–4, e.g. identify, calculate, sequence, contrast, identify a relationship, compare, infer, predict and draw a conclusion
  - were of appropriate scope and scale
  - used the number of marks to indicate the number of cognitive processes required in the response

- used appropriate cognitive verbs aligned to the nature of the expected response, e.g. contrast questions require recognition of differences by deliberate juxtaposition of contrary elements (Syllabus glossary).

Practices to strengthen

It is recommended that assessment instruments:

- align questions to the correct objective (see the Mark allocations table, Syllabus section 4.5.1 (AS section 2.5.1)), e.g. questions requiring students to
  - determine unknown scientific quantities or features align to Objective 2 (apply understanding)
  - identify trends, patterns, relationships, limitations or uncertainty in datasets align to Objective 3 (analyse evidence)
  - draw conclusions based on the analysis of datasets align to Objective 4 (interpret evidence).
- avoid use of part marks as this can affect the scale and transparency of the item, i.e. students may use the number of marks when planning their response.

Accessibility

Accessibility in assessment design ensures that no student or group of students is disadvantaged in their capacity to access an assessment.

Reasons for non-endorsement by priority of assessment

Accessibility priority	Number of times priority was identified in decisions*
Bias avoidance	32
Language	88
Layout	42
Transparency	48

*Each priority might contain up to four assessment practices.

Total number of submissions: 439.

Effective practices

Accessibility priorities were effectively demonstrated in assessment instruments that:

- contained succinct questions/commands that clearly cued the expected responses identified in the marking scheme, e.g. by ensuring the cognitive verb matched the nature of the response (refer to Syllabus glossary)
- used images that were clear, relevant and free from distractors
- were free from spelling, grammar and punctuation errors
- had clearly labelled figures and graphs.

Practices to strengthen

It is recommended that assessment instruments:

- use correct scientific binomial nomenclature for species names, e.g. *Monodon monoceros/ M. monoceros.*

Additional advice

- Schools should
  - use the ‘Print preview’ function to ensure that graphs and/or images in the datasets are appropriately placed and not split over two pages
  - ensure internal quality assurance processes are carried out before submitting instruments for endorsement (*QCE and QCIA policy and procedures handbook v5.0,* Section 9.6.1*)*.

Assessment decisions

Reliability

Reliability is a judgment about the measurements of assessment. It refers to the extent to which the results of assessments are consistent, replicable and free from error.

Agreement trends between provisional and confirmed marks

Criterion number	Criterion name	Percentage agreement with provisional	Percentage less than provisional	Percentage greater than provisional	Percentage both less and greater than provisional
1	Data test	95.85%	3%	0.92%	0.23%

Effective practices

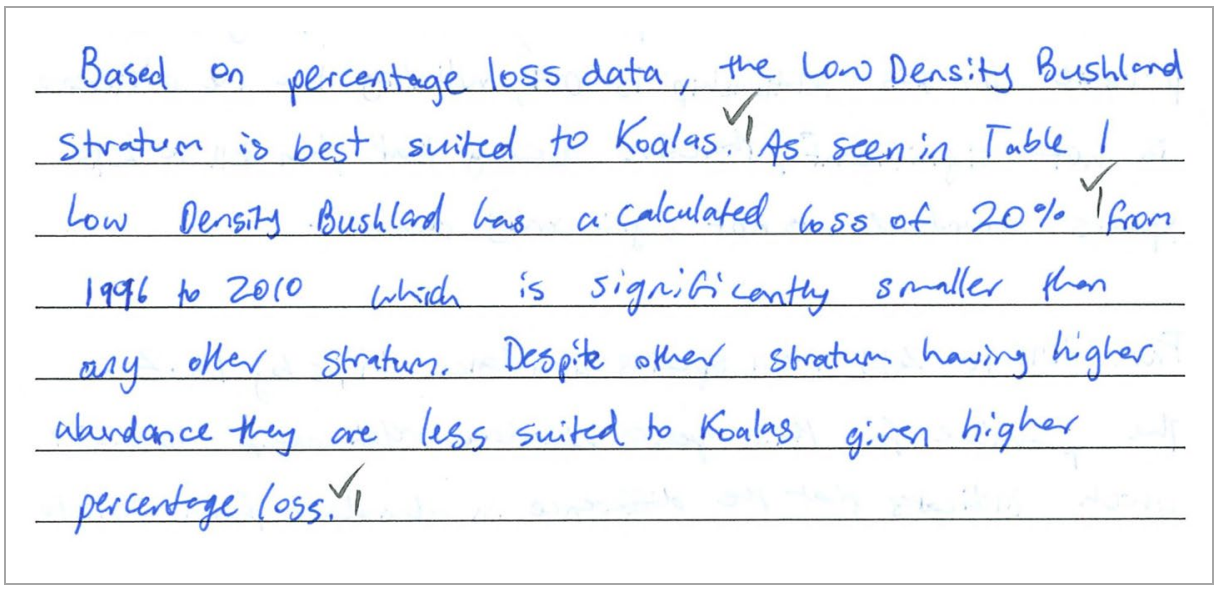
Accuracy and consistency of the application of the ISMG for this IA was most effective when:

- the marking scheme demonstrated the mark allocations for each aspect of the response awarded a mark
- alternative responses were captured in the marking scheme and annotated to show how marks were allocated
- percentage cut-offs were applied appropriately to determine a provisional mark.

Samples of effective practices

The following excerpt has been included to demonstrate the use of annotations on a student response to an Objective 4 (interpret evidence) item that required a conclusion be drawn based on analysis of a dataset.

Note: The characteristic/s identified may not be the only time the characteristic/s has occurred throughout a response.


Based on percentage loss data, the Low Density Bushland Stratum is best suited to Koalas. ✓1 As seen in Table 1 Low Density Bushland has a calculated loss of 20% ✓1 from 1996 to 2010 which is significantly smaller than any other stratum. Despite other stratum having higher abundance they are less suited to Koalas given higher percentage loss. ✓1

Practices to strengthen

To further ensure accuracy and consistency of the application of the ISMG for this IA, it is recommended that:

- teachers annotate the student response for each question to clearly indicate how marks were awarded in accordance with the marking scheme (*QCE and QCIA policy and procedures handbook* v5.0, Section 9.7.1).

Additional advice

- Schools are required to submit samples of student assessment responses for review during confirmation. Submissions should align with the relevant confirmation information (*QCE & QCIA policy and procedures handbook v5.0*, Section 9.7.1). Scanned student samples should be checked to ensure they are complete and legible before being submitted for confirmation. The *Confirmation submission information* for Biology is available under Resources in the Syllabuses application (app) on the QCAA Portal.
- Comparable assessments should be developed in the Endorsement app to ensure the correct examination and its matching marking scheme are available for confirmation (*QCE and QCIA policy and procedures handbook v5.0,* Section 7.4).

Internal assessment 2 (IA2)

Student experiment (20%)

This assessment requires students to research a question or hypothesis through collection, analysis and synthesis of primary data. A student experiment uses investigative practices to assess a range of cognitions in a particular context. Investigative practices include locating and using information beyond students' own knowledge and the data they have been given.

Research conventions must be adhered to. This assessment occurs over an extended and defined period of time. Students may use class time and their own time to develop a response.

Assessment design

Validity

Validity in assessment design considers the extent to which an assessment item accurately measures what it is intended to measure and that the evidence of student learning collected from an assessment can be legitimately used for the purpose specified in the syllabus.

Reasons for non-endorsement by priority of assessment

Validity priority	Number of times priority was identified in decisions*
Alignment	15
Authentication	15
Authenticity	3
Item construction	12
Scope and scale	1

*Each priority might contain up to four assessment practices.

Total number of submissions: 439.

Effective practices

Validity priorities were effectively demonstrated in assessment instruments that:

- included all task requirements as part of the task description (Syllabus section 4.5.2, AS section 2.5.2)
- provided appropriate checkpoints and scaffolding consistent with the advice in the *QCE and QCIA policy and procedures handbook v5.0*, Section 8.

Practices to strengthen

It is recommended that assessment instruments:

- include authentication strategies to indicate how group work will be monitored, e.g. 'the teacher will compare the responses of students who have worked together'
- ensure that the practicals listed in the task align with the Unit 3 topics selected in the conditions section
- ensure scaffolding does not lead students to a predetermined response by identifying what information should be included in each paragraph or section.

Accessibility

Accessibility in assessment design ensures that no student or group of students is disadvantaged in their capacity to access an assessment.

Reasons for non-endorsement by priority of assessment

Accessibility priority	Number of times priority was identified in decisions*
Bias avoidance	0
Language	6
Layout	4
Transparency	4

*Each priority might contain up to four assessment practices.

Total number of submissions: 439.

Effective practices

Accessibility priorities were effectively demonstrated in assessment instruments that:

- avoided unnecessary repetition of the task in different sections, e.g. task specifications were only provided in the task description
- were free from errors and modelled accurate spelling, grammar, punctuation and other textual features
- used correct scientific binomial nomenclature for species names, e.g. *Monodon monoceros/ M. monoceros*.

Practices to strengthen

There were no significant issues identified for improvement.

Assessment decisions

Reliability

Reliability is a judgment about the measurements of assessment. It refers to the extent to which the results of assessments are consistent, replicable and free from error.

Agreement trends between provisional and confirmed marks

Criterion number	Criterion name	Percentage agreement with provisional	Percentage less than provisional	Percentage greater than provisional	Percentage both less and greater than provisional
1	Research and planning	90.53%	9.24%	0.23%	0%
2	Analysis of evidence	92.61%	6.93%	0.46%	0%
3	Interpretation and evaluation	92.15%	7.85%	0%	0%
4	Communication	99.08%	0%	0.92%	0%

Effective practices

Accuracy and consistency of the application of the ISMG for this IA was most effective when:

- for the Interpretation and evaluation criterion,
  - *justified* discussion of reliability used reasoning and evidence (e.g. patterns in uncertainty data) to consider how specific aspects of the experimental design or data collection process affected the extent to which another experimenter would obtain the same results
  - *justified* discussion of validity had clear links to the purpose of the experiment (e.g. as stated in the rationale), and considered how specific aspects of the experimental design affected the extent to which the experiment measured what was intended
  - *logically derived* improvements addressed aspects of the experiment identified as affecting reliability, e.g. introducing counting criteria to improve precision of quadrat data
  - *logically derived* extensions addressed aspects of the experiment identified as affecting validity, e.g. increasing the range of pHs an enzyme is exposed to when expected tolerance curves are not obtained (AS Unit 1).
- for the Communication criterion,
  - understanding of the subject matter and experimental findings were effectively conveyed using *fluent* and *concise* scientific language and representations, e.g.
    - subject- and discipline-specific terms such as species names, 'reliability' and 'validity' were used correctly
    - numerical data was expressed to an appropriate number of decimal places and, where relevant, using scientific notation
    - appropriate units and symbols were used throughout the response
  - findings were communicated using *appropriate* genre conventions, e.g. for scientific reports
    - tables and graphs were appropriately titled/captioned and labelled
    - statistical measures represented by error bars were clearly stated in the graph title or caption
    - findings were presented in appropriate sections.

Samples of effective practices

The following excerpts demonstrate a rationale that considers AS Unit 1 subject matter (i.e. factors affecting the reaction rate of enzymes), leading to a specific and relevant research question.

Note: The characteristic/s identified may not be the only time the characteristic/s has occurred throughout a response.

Excerpt 1

Enzymes are biological catalysts made from proteins that are in all living things, lowering activation energy to speed up biochemical reactions and support cellular metabolism (*Figure 1*). Studying enzymes can reveal diverse human body reactions, aiding our comprehension and prediction of metabolism in living organisms. (Jack Davis, 2019).

Excerpt 2

Enzymes bind to specific substrates via their active site to form an enzyme-substrate complex (*Figure 2*). Catalase's substrate is H_2O_2 (hydrogen peroxide) which is decomposed into water (H_2O) and oxygen gas (O_2) through a neutralisation reaction (*Figure 3)* (Khan Academy, 2023). The volume of O_2 (product of reaction) displaced during this reaction serves as a measure for the effectiveness of the catalase's breakdown of H_2O_2. This experiment will employ a 3% H_2O_2 solution, as in the original experiment.

Factors influencing the efficiency of the reaction include temperature, pH, and enzyme/substrate concentrations. Due to the sensitive nature of the enzyme-substrate complex, enzymes have specific optimal conditions for optimal activity with deviations resulting in denaturation (inactivity due to structural changes) (Zheng, 2018). The impact of pH on the activity of the catalase enzyme found in potato will be explored to determine the optimal pH level.

The optimal pH level for catalase varies depending on the vegetable or tissue due the pH of its environment, with potato-derived catalase functioning best at pH 9 (UKEssays, 2018). Thus, the modified experiment's hypothesis is that potato-derived catalase will displace the most O_2 and perform optimally at pH 9 when reacted with the 3% H_2O_2 solution. It is predicted that above this pH, denaturation will occur, significantly reducing O_2 volume production as the enzyme cannot bind to its substrate (H_2O_2) due to structural changes.

Excerpt 3

Research question

To what extent does the pH of a solution effect the activity of the enzyme catalase in potato when placed in 3% H_2O_2 solution, measured through the volume (mL) of O_2 displaced?

The following excerpts demonstrate justified modifications to the methodology and considered management of risks and environmental issues for a Unit 3 (General syllabus) research question.

Note: The characteristic/s identified may not be the only time the characteristic/s has occurred throughout a response

Excerpt 1

"How does increasing the concentration of air pollutants, demonstrated by PM2.5 concentration (μg/m³), influence the percentage cover of green foliose lichen on yellow poinciana trees (P.pterocarpum) at various locations within Cairns?"

Excerpt 2

- *Experiment redirection from impact of orientation on lichen percentage cover, to the impact of PM2.5:* pollution is a resoundingly negative abiotic factor, with resulting bleaching and decreased photosynthesis (Sett & Kundu 2016). Quantifying its impact helps inform local environmental decisions to protect lichen diversity.
- *Ten locations chosen:* the following sites were chosen within proximity to roads, dumps, and rainforests to vary PM2.5 concentration (independent variable) to measure the impact on green foliose lichen percentage cover (dependent variable). Having ten variations of the independent variable decreases the effect of an outlier on the overall trend, thereby increasing the trend's reliability (Van Belle 2008).
    - Yorkey's Knob: dump, Varley Street.
    - Smithfield: roundabout, shopping centre carpark, fuel station.
    - Redlynch State College: industrial technology block, multi-purpose centre.
    - Redlynch: Lum Jim Street, Redlynch Park, Goomboora park.
- *Trials increased from 1 to 5:* increasing tree sample sizes increases the likelihood mean lichen percentage cover will be accurate to true values, hence will increase the validity of results. Larger sample size also allows Inter-Quartile-Range to be calculated, increasing the reliability of data by allowing potential outliers to be removed systematically (Charlesworth Author Services 2022).
- *Counting criteria introduced:* visual estimation can be both imprecise and inaccurate hence a counting criteria was used so lichen percentage cover would be more precise, and therefore more reliable (Rosso, Neitlich & Smith 2014). Since if lichen is present in any part of 1 of 100 squares it is counted as a whole percent, this will overestimate lichen cover. However, visual estimation can both under and overestimate percentage cover. Therefore, the counting criteria will be more reliable.
- *Surveyed point on made 30cm from the ground at the eastern point:* east was chosen since sampling was done in the morning, so lichens on the east side are clearly illuminated, increasing observation accuracy, and therefore validity of lichen percentage covers. Height affects lichen growth, hence was kept at 30cm from a consistent point (east) to keep heights uniform which ensures this extraneous variable is standardised, limiting random error, and protecting reliability of results (Wietrzyk, Węgrzyn & Lisowska 2016).

Excerpt 3

- *Potential damage to microhabitats on P.pterocarpum with quadrat:* to minimise this, selected lichens were foliose which have a flat thallus, therefore are more difficult to dislodge and cause harm to (Smith 2014).
- *Danger of accidents due to sampling next to busy road:* to limit risk, high-visibility clothing was worn.
- *Potential toxicity of tree sap:* gloves were worn to protect skin, as were sunglasses to protect eyes.

The following excerpt demonstrates correct and relevant processing of data and thorough identification of trends, patterns, relationships and uncertainty for a Unit 3 research question.

Note: The characteristic/s identified may not be the only time the characteristic/s has occurred throughout a response.

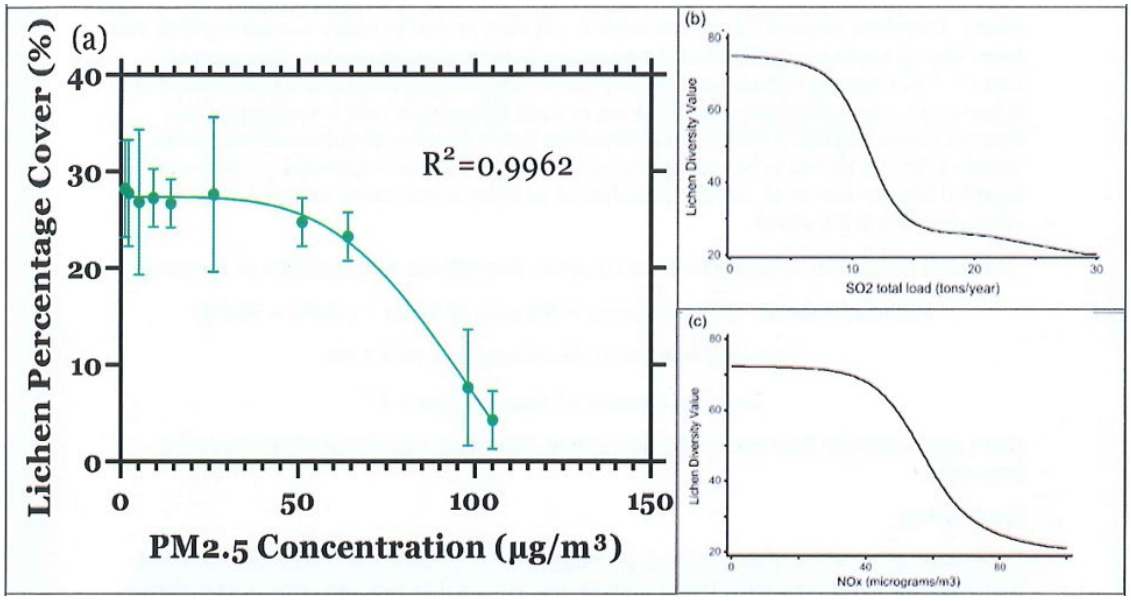


Figure 2. Impact of increased PM2.5 Concentration on Lichen Percentage Cover within Cairns (2a), and Giordani's (2006) research (2b/2c).

Figure 2a supports Giordani's (2006) research (Figure 2b/2c): when PM2.5 concentration increases, green foliose lichen percentage cover decreases, in a negative logistic trend. A logistic trend is appropriate since it can fit within 0-100% for infinite PM2.5 concentrations and is in-keeping with Giordani's (2006) research (Figure 2b/2c). R^2 values denote how well the logistic regression describes changes in lichen percentage cover for collected data: an R^2 value of 0.9962 indicates high correlation, since it is close to 1 (Jin et al. 2001). Initial results (1-64µg/m³PM2.5) indicated stable lichen percentage cover of 26.67-28.20%, characterised by all error bars overlapping, meaning any differences weren't statistically significant. However, a limit for PM2.5 concentration tolerance in lichen was reached at approximately 64µg/m³PM2.5, denoted by a large, and statistically significant drop in mean lichen percentage cover (Figure 2a) from 23.25% at 64µg/m³ to 7.6% at 98µg/m³. Regardless of PM2.5 concentration, absolute uncertainty fluctuates between 2.5-7.5%, indicating uncertainty is unrelated to PM2.5 concentration, and was possibly affected by other abiotic factors which affect lichen growth (Renhorn et al. 1997).

Practices to strengthen

To further ensure accuracy and consistency of the application of the ISMG for this IA, it is recommended that:

- for the Research and planning criterion,
  - a *considered* rationale clearly connects the research question to Unit 3 (or AS Unit 1) subject matter, establishes a logical basis for the experiment, and explains the relevance of the independent and dependent variables

  - a *relevant* research question for Unit 3 allows students to investigate biodiversity or ecosystem dynamics, e.g. when abiotic factors are investigated, it should be in the context of their effects on biodiversity or ecosystem dynamics
  - *justified* modifications to the methodology clearly state how each modification will improve the reliability or validity of the evidence
  - the impacts of risks associated with the experiment and their subsequent management are *considered* with regard to how the methodology was carried out.
- for the Analysis of evidence criterion,
  - *correct* and *relevant* processing of data shows discriminating selection and accurate use of techniques, summary statistics and graphical representations appropriate to the research question, e.g.
    - line graphs with error bars to show trends
    - scatterplots with correlation coefficients to identify relationships
    - bar graphs with error bars and Student's t-tests to make comparisons
    - standard deviation and/or standard error to identify uncertainty
  - *thorough* identification of limitations focuses on aspects of the data that make the evidence less effective for the purpose of responding to the research question.

Additional advice

- When determining best-fit on an ISMG, the higher mark in the performance level should only be awarded if there is evidence of all the characteristics in the performance-level descriptor or better (see *Using ISMGs for General Science syllabuses* in the Syllabuses app on the QCAA Portal). Marked ISMGs should indicate the characteristics evident in the student response and the mark awarded for each criterion (*QCE and QCIA policy and procedures handbook v5.0*, Section 9.7.1).
- Student experiments using databases as a source of data must include
  - evidence that indicates how students engaged with the database and how they refined, redirected or extended the investigation through collection of primary data
  - primary data collected by the student.
- Unique responses to the assessment task must be developed, particularly when students work in groups to plan and conduct the experiment and collected identical data. Further information about authenticating student responses can be found in the *QCE and QCIA policy and procedures handbook v5.0*, Sections 8.2.8 and 11.1.5.
- Opportunities to demonstrate the relevant data-processing techniques that can be used to identify trends/patterns/relationships and uncertainty/limitations of experiment data should be embedded as part of the teaching and learning. This supports students to draw meaning from the data they collect in their experiments

Internal assessment 3 (IA3)

Research investigation (20%)

This assessment requires students to evaluate a claim. They will do this by researching, analysing and interpreting secondary evidence from scientific texts to form the basis for a justified conclusion about the claim. A research investigation uses research practices to assess a range of cognitions in a particular context. Research practices include locating and using information beyond students' own knowledge and the data they have been given.

Research conventions must be adhered to. This assessment occurs over an extended and defined period of time. Students may use class time and their own time to develop a response.

Assessment design

Validity

Validity in assessment design considers the extent to which an assessment item accurately measures what it is intended to measure and that the evidence of student learning collected from an assessment can be legitimately used for the purpose specified in the syllabus.

Reasons for non-endorsement by priority of assessment

Validity priority	Number of times priority was identified in decisions*
Alignment	14
Authentication	11
Authenticity	2
Item construction	7
Scope and scale	0

*Each priority might contain up to four assessment practices.

Total number of submissions: 435.

Effective practices

Validity priorities were effectively demonstrated in assessment instruments that:

- provided claims with clear links to Unit 4 (or AS Unit 2) subject matter and avoided redirection into ethical issues that may not allow students to effectively complete all elements of the task
- included all task requirements in the task description (Syllabus section 5.5.1, AS section 3.5.1)
- listed appropriate authentication strategies to ensure unique student responses.

Practices to strengthen

It is recommended that assessment instruments:

- avoid including authentication strategies that suggest aspects of the task are to be completed in groups, as this contradicts the task conditions outlined in Syllabus section 5.5.1 (AS section 3.5.1).

Accessibility

Accessibility in assessment design ensures that no student or group of students is disadvantaged in their capacity to access an assessment.

Reasons for non-endorsement by priority of assessment.

Accessibility priority	Number of times priority was identified in decisions*
Bias avoidance	1
Language	13
Layout	3
Transparency	0

*Each priority might contain up to four assessment practices.

Total number of submissions: 435.

Effective practices

Accessibility priorities were effectively demonstrated in assessment instruments that:

- were free from spelling, punctuation and grammatical errors.

Practices to strengthen

There were no significant issues identified for improvement.

Assessment decisions

Reliability

Reliability is a judgment about the measurements of assessment. It refers to the extent to which the results of assessments are consistent, replicable and free from error.

Agreement trends between provisional and confirmed marks

Criterion number	Criterion name	Percentage agreement with provisional	Percentage less than provisional	Percentage greater than provisional	Percentage both less and greater than provisional
1	Research and planning	93.3%	5.08%	1.39%	0.23%
2	Analysis and interpretation	93.3%	5.54%	0.92%	0.23%
3	Conclusion and evaluation	90.76%	8.78%	0.23%	0.23%
4	Communication	98.85%	0%	1.15%	0%

Effective practices

Accuracy and consistency of the application of the ISMG for this IA was most effective when:

- for the Analysis and interpretation criterion,
  - enough secondary evidence was included to validly address all aspects of the research question, e.g. multiple sources of evidence allowed limitations of one source to be addressed using another
  - *thorough* and appropriate identification of limitations was specific to the research question and identified how or why each limitation made the evidence less effective, e.g. considering
    - the amount of reliable and current research
    - sample size and relevance in the academic studies
    - the validity of the testing processes used in the studies
    - which aspect/s of the research question each piece of evidence addressed
  - *justified* scientific arguments were supported by evidence, linked to the research question, and demonstrated understanding of the relationship between the different sources of evidence in answering the research question.
- for the Communication criterion,
  - understanding of the subject matter and research findings were effectively conveyed using appropriate scientific language and representations, e.g. by accurately using subject-specific terminology, units and symbols throughout the report
  - appropriate use of referencing conventions was evident through consistent use of in-text referencing and a complete bibliography or reference list.

Samples of effective practices

The following excerpt demonstrates a considered rationale, showing clear development of the research question from the claim. The student was responding to a claim relating to the use of genetic testing to predict sporting performance.

Note: The characteristic/s identified may not be the only time the characteristic/s has occurred throughout a response.

(Franceschini, Frick Kopp, 2018). However, genetic screening has also shown importance within the sports industry through its ability to identify genes that are linked to sporting ability within sport: swimming, mountaineering, rowing, and running athletes. Over recent years the popularity of this for talent identification has skyrocketed with a known 70 companies offering direct-to-consumer genetic testing that assist in the prediction of sports performance/injury (Pickering, Kiely, Grgic, Lucia, Coso, 2019).

Over 200 genetic markers have been tentatively linked with physical performance (Varillas-Delgado et al, 2022). Some of these include AMPD1, ACTN3, CYP2D6, BCDIN3D and PPARGC1A, studies have shown that these genes play a key role in sports performance through their related phenotypes, of power, strength, aerobic capacity, flexibility and temperament (Ahmetov, Egorova, Gabdrakhmanova, Fedotovskaya, 2016).

The angiotensin I-converting enzyme (ACE) is frequently studied due to its important role within the renin-angiotensin-aldosterone system (RAS), influencing homeostasis (Jones, Montgomery, Woods, 2002). ACE was the first genetic marker linked to sports performance, it is located in the 287-bp fragment in intron 16 and contains polymorphisms consisting of the insertion (I) and/or deletion (D) allele/s (DD, ID, II) (Panahloo et al, 1995 & Guth, Roth, 2013). Through studies on various elite athletes, it is suggested that each allele influences the athlete's physical abilities potentially increasing their performance in their respective sport. These studies demonstrated the *D* allele is associated with strength and power, while the I allele is connected to endurance events.

ACE catalyses the conversion of the hormone angiotensin I into angiotensin II (Varillas-Delgado et al, 2022). This is more prevalent within the D allele as it obtains a higher ACE activity, therefore, producing more angiotensin II (Eider et al, 2013).

Angiotensin II, degrades bradykinin, a small protein that induces vasodilation and vascular permeability, by breaking it down the body is able to control blood pressure and inflammation throughout the body (Woods, Humphries, Montgomery, 2000). Additionally, The D allele is associated with a greater proportion of fast contraction muscle fibres, while the I allele is linked to a higher concentration of slow-twitch muscle fibres (Yang, Lin, Jia, Chen, 2023). The factors associated with the D allele in ACE is advantageous for power athletes like sprinters as, regulating blood pressure controls hypertension, resulting in higher VO2 max, meaning the body can absorb more oxygen, utilising it during athletic performance (Sawada, Tanaka, Funakoshi, Shindo, Kono, Ishiko, 1993 & Scribbans, Vecsey, Hankinson, Foster, Gurd, 2016). Additionally, fast twitch muscle fibres are utilised for rapid movements, contracting quickly and forcefully using an anaerobic process, which is ideal for sprinting as they are not limited by need for oxygen intake (LibreTexts Medicine, 2023).

If research suggests that there is a correlation between D allele genotypes and elite sprint swimming athletes in their phenotypes, then this would support the claim. Therefore, the research question will focus on whether the D/D polymorphism of the ACE gene influences athletes' performance in sprints events.

Is there a significant difference in the allele frequency distribution (P≤0.05) of the D allele ACE polymorphism in elite sprint swimming athletes for events (events ≤400m)?

Samples of effective practices

The following excerpts demonstrate justified conclusions linked to the research question and a suggested extension to the investigation that is considered and relevant to the claim.

Note: The characteristic/s identified may not be the only time the characteristic/s has occurred throughout a response.

Excerpt 1

Does the gene therapy, Valoctocogene roxaparvovec [redacted], administered through the AAV5 vector, have the potential to decrease the annualised bleeding rate of male Haemophilia type A patients by an average of at least 72% over 3 years?

Both studies display a decreasing trend for the annualised bleeding rate of patients with Haemophilia A after the administration of [redacted] gene therapy using the AAV5 vector. Figure 1 indicates that the annualised bleeding rate can be reduced by more than 72% with an average decline in bleeding of 89.17% over 4 years (17.17% greater than benchmark). In corroboration with such a decline, Figure 2's patients bleeding rates decreased by 96% over 3 years (24% greater than benchmark). However, various limitations impact the conclusions such as limited sample size, which limits the ability for extrapolating the data to all populations and variations in dosages and vector infusions, which decreases the consistency of the studies and leads to various patients not experiencing effective treatment. Despite such factors, it is predicted that gene therapy is effective in reducing the average bleeding rate of patients by 72% over 3 years.

In relation to the claim, *'gene therapy is an effective way to treat diseases',* the research collected suggests that gene therapy is an effective way to treat disease. The research question used the gene therapy of [redacted] to treat the disease Haemophilia A. Regarding 'effective treatment', the research question specified an average reduction in annualised bleeding rate of at least 72% for 3 years. This was met by both studies with average reductions of 89.17% and 96% over at least 3 years for both. Therefore, the claim is supported as gene therapy was deemed effective when treating Haemophilia A. However, this investigation only studied patients for 3 years and for only one type of gene therapy hence, further study is recommended to extrapolate to the claim more conclusively.

Excerpt 2

Investigate the use of gene therapy in treating other diseases such as the commonly studied disease, cystic fibrosis (Diseases Treated by Gene Therapy, 2023). Research states that gene therapy within cystic fibrosis is safe and results in small improvements in lung function (Gene Therapy for Cystic Fibrosis , 2023) therefore, although gene therapy is effective for Haemophilia A, its effectiveness or safety may be compromised when treating other diseases. Therefore, extending the investigation to other diseases would enable a more conclusive and valid answer to the claim.

Practices to strengthen

To further ensure accuracy and consistency of the application of the ISMG for this IA, it is recommended that:

- for the Research and planning criterion,
  - a *considered* rationale integrates understanding of Unit 4 (or AS Unit 2) concepts and information from credible sources to clearly show how the research question was developed from the claim
  - a *specific* and relevant research question is clearly defined, logically developed through the rationale and allows for effective investigation of Unit 4 (or AS Unit 2) topics

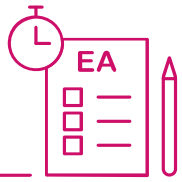
  - *sufficient* sources provide enough scope to develop the research question and enough secondary evidence to draw valid conclusions that can be extrapolated to the claim

- for the Conclusion and evaluation criterion,
  - *insightful discussion* of the quality of evidence considers how well the evidence answers the research question and the extent to which limitations identified in the analysis of evidence influence the validity of the conclusion
  - *justified* conclusions are backed by evidence, specific to the research question, and based on scientific arguments developed throughout the response
  - improvements and extensions focus on ways the research investigation could be refined or extended to obtain more valid evidence applicable to the claim.

Additional advice

- A variety of sources should be used to enhance understanding of the claim, develop the research question and provide sufficient evidence. Students should continually reflect on the quality of evidence throughout the research process to ensure appropriate evidence is obtained. For further information, see *IA3 effective processes and practices: Selecting sources* under Resources in the Syllabuses app on the QCAA Portal.

External assessment



External assessment (EA) is developed and marked by the QCAA. The external assessment for a subject is common to all schools and administered under the same conditions, at the same time, on the same day.

Examination (50%)

Assessment design

The assessment instrument was designed using the specifications, conditions and assessment objectives described in the summative external assessment section of the syllabus.
The examination consisted of two papers:

- Paper 1, Section 1 consisted of multiple choice questions (20 marks)
- Paper 1, Section 2 consisted of short response questions (26 marks)
- Paper 2, Section 1 consisted of short response questions (43 marks)

The examination assessed subject matter from Units 3 and 4. Questions were derived from the contexts of Describing biodiversity, Ecosystem dynamics, DNA genes and the continuity of life, Continuity of life on Earth.

The assessment required students to respond to multiple choice and short response questions.

The AS assessment instrument was designed using the specifications, conditions and assessment objectives described in the summative external assessment section of the AS.
The AS examination consisted of two papers:

- Paper 1, Section 1 consisted of multiple choice questions (20 marks)
- Paper 1, Section 2 consisted of short response questions (25 marks)
- Paper 2, Section 1 consisted of short response questions (46 marks)

The AS examination assessed subject matter from AS Units 1 and 2. Questions were derived from the contexts of Cells as the basis of life, Multicellular organisms, Homeostasis and Infectious disease.

The AS assessment required students to respond to multiple choice and short response questions.

Assessment decisions

Assessment decisions are made by markers by matching student responses to the external assessment marking guide (EAMG). The external assessment papers and the EAMG are published in the year after they are administered.

Multiple choice question responses

There were 20 multiple choice questions in Paper 1.

Percentage of student responses to each option

Note:

- The correct answer is **bold** and in a blue shaded table cell.
- Some students may not have responded to every question.

Biology General: Paper 1

Question	A	B	C	D
1	11.49	13.37	3.57	**71.45**
2	2.59	11.74	**80.95**	4.61
3	9.65	**71.31**	11.21	7.62
4	1.55	9.55	**80.9**	7.94
5	4.09	13.16	15.08	**67.58**
6	**73.37**	7.52	8.45	10.51
7	8.98	3.44	8.14	**79.34**
8	9.6	5.75	**68.78**	15.78
9	6.89	**76.21**	12.78	4.05
10	**97.89**	0.44	0.58	1.02
11	2.38	**76.57**	5.21	15.73
12	5.26	4.74	**81.52**	8.39
13	18.15	23.2	**40.63**	17.74
14	5.69	7.62	6.05	**80.52**
15	**32.26**	57.72	6.42	3.45
16	**54.43**	18.67	17.38	9.38
17	8.27	**42.67**	28.57	20.33
18	**76.05**	6.23	6.46	11.14
19	9	**46.28**	29.94	14.57
20	3.25	5.86	13.22	**77.49**

Biology AS: Paper 1

Question	A	B	C	D
1	**80.87**	3.32	8.93	6.38
2	**42.35**	8.16	20.92	27.81
3	3.57	29.85	**52.81**	13.27
4	12.24	6.12	20.15	**60.97**
5	**16.84**	19.39	26.02	37.24
6	19.9	5.87	**59.18**	14.54
7	13.27	15.82	17.09	**53.32**
8	16.07	22.96	**45.66**	14.29
9	9.44	19.9	5.1	**65.05**

Question	A	B	C	D
10	36.48	28.83	**23.98**	10.2
11	18.88	**49.23**	12.24	19.13
12	11.99	**47.7**	5.36	34.44
13	**63.27**	6.12	23.72	6.38
14	**41.84**	33.67	12.5	11.48
15	24.74	30.36	4.34	**40.05**
16	9.95	3.57	66.84	**19.13**
17	7.14	**82.65**	6.12	3.32
18	2.04	9.44	32.4	**54.85**
19	**44.13**	2.3	13.52	38.78
20	21.43	**53.83**	21.17	2.81

Effective practices

Overall, students responded well to:

- questions requiring analysis and interpretation of data from a stimulus, e.g.
  - analysing a food web to determine the effect of removing a keystone species
  - inferring phylogenetic relationships from molecular sequence data
  - interpreting field data to classify an ecosystem
- questions requiring identification of an unknown value from a graph or stimulus.

Samples of effective practices

Short response

The following excerpt is from Question 25c) in Paper 1. It required students to explain why the provided data on species richness and Simpson's diversity index (SDI) was more informative than a single measure when investigating the effect of an invasive species on plant biodiversity.

Effective student responses:

- recognised species richness is the total number of species present
- recognised SDI considers species number and abundance/evenness
- used the data to explain why data on species richness and SDI is more informative than a single measure.

This excerpt has been included:

- to demonstrate an explanation supported by data from the stimulus.

Species richness is a measure of how many different species are present at the ecosystem, while Simpson's diversity index assesses biodiversity and is the probability that two randomly selected organisms are of different species. Both of these measures are important for assessing biodiversity, as together they can provide more comprehensive information on the number of different species and species evenness, where a single measure cannot. For example, species richness remained constant at 7 for the invasive species cover of 0-60%, but SDI decreases from 0.83 to 0.55, revealing that while there is the same number of species, the species evenness is likely becoming more skewed, thus biodiversity is decreasing. On the other hand, SDI remained relatively constant at 0.55 and 0.49 for 40-60% and 60-80% cover respectively, but species richness decreased from 7 to 4, revealing that the number of species is decreasing, supporting lower biodiversity. These reasons evidence how using both measures provide more comprehensive information on biodiversity than just one.

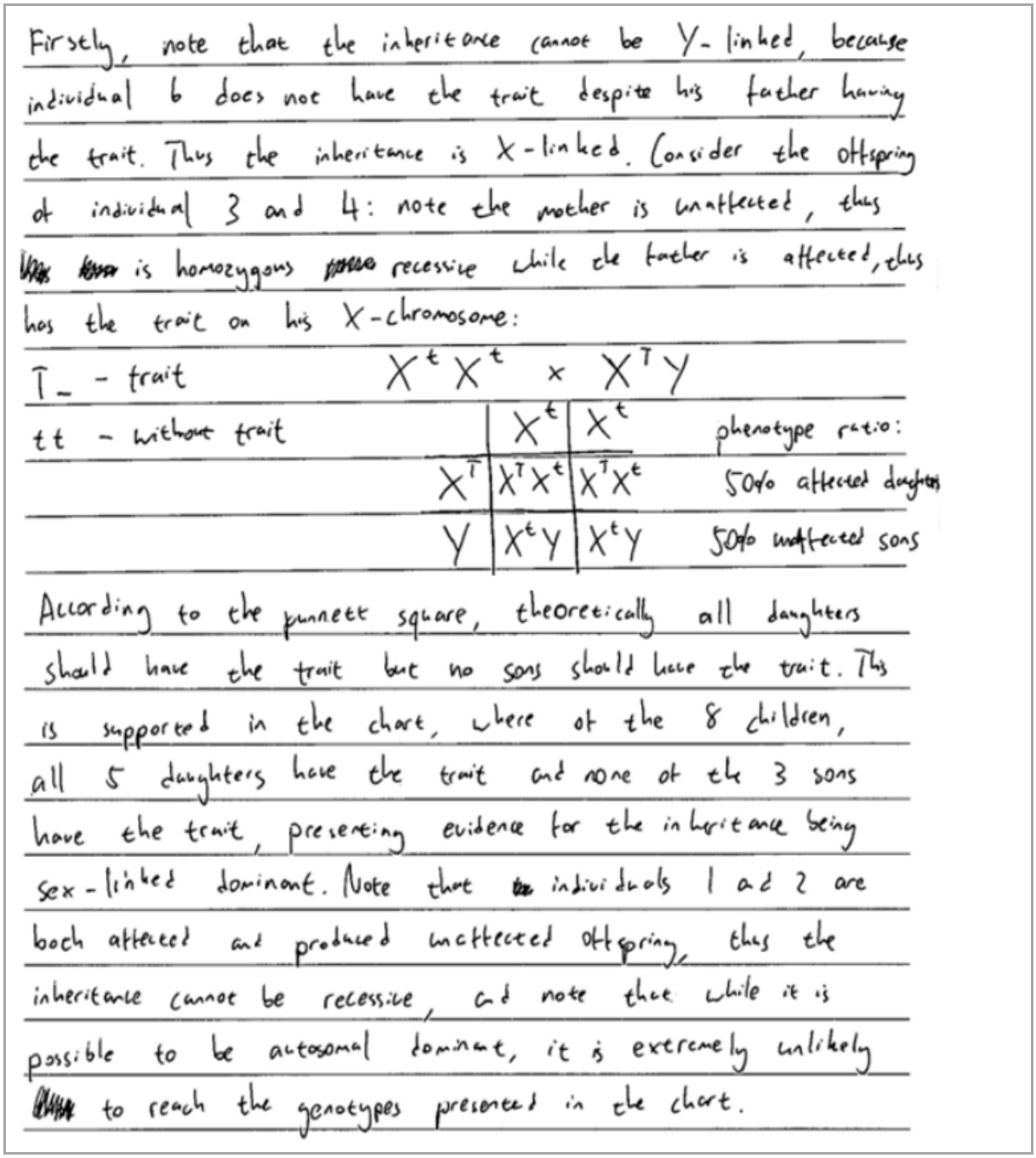
The following excerpt is from Question 3b) in Paper 2. It required students to identify evidence that a trait is sex-linked dominant from a pedigree and justify their response using a Punnett square.

Effective student responses:

- identified appropriate evidence from the pedigree
- identified appropriate genotypes for the Punnett square
- constructed an appropriate Punnett square
- used the Punnett square to determine expected frequencies for offspring phenotypes.

This excerpt has been included:

- to demonstrate how the student used evidence from the stimulus to develop a response.



Firstly, note that the inheritance cannot be Y-linked, because individual 6 does not have the trait despite his father having the trait. Thus the inheritance is X-linked. Consider the offspring of individual 3 and 4: note the mother is unaffected, thus is homozygous recessive while the father is affected, thus has the trait on his X-chromosome:

T_ - trait

tt - without trait

$X^tX^t \times X^TY$

	X^t	X^t	phenotype ratio:
X^T	X^TX^t	X^TX^t	50% affected daughters
Y	X^tY	X^tY	50% unaffected sons

According to the punnett square, theoretically all daughters should have the trait but no sons should have the trait. This is supported in the chart, where of the 8 children, all 5 daughters have the trait and none of the 3 sons have the trait, presenting evidence for the inheritance being sex-linked dominant. Note that individuals 1 and 2 are both affected and produced unaffected offspring, thus the inheritance cannot be recessive, and note that while it is possible to be autosomal dominant, it is extremely unlikely to reach the genotypes presented in the chart.

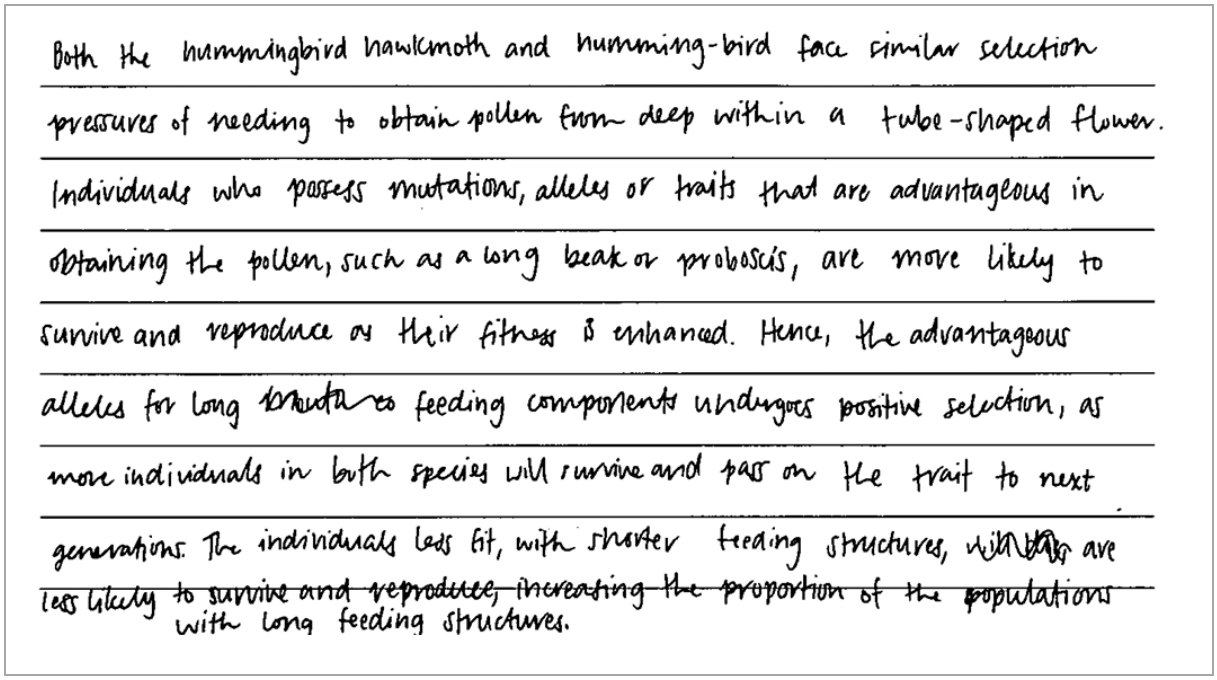
The following excerpt is from Question 24b) in Paper 2. It required students to use the principles of natural selection to explain similarities between two species.

Effective student responses:

- recognised the two species are exposed to similar selection pressures
- explained how natural selection favours similar features.

This excerpt has been included:

- to demonstrate a response that refers to the principles of natural selection.


Both the hummingbird hawkmoth and humming-bird face similar selection pressures of needing to obtain pollen from deep within a tube-shaped flower. Individuals who possess mutations, alleles or traits that are advantageous in obtaining the pollen, such as a long beak or proboscis, are more likely to survive and reproduce as their fitness is enhanced. Hence, the advantageous alleles for long ~~breath~~es feeding components undergoes positive selection, as more individuals in both species will survive and pass on the trait to next generations. The individuals less fit, with shorter feeding structures, will ~~be~~ are less likely to survive and reproduce, increasing the proportion of the populations with long feeding structures.

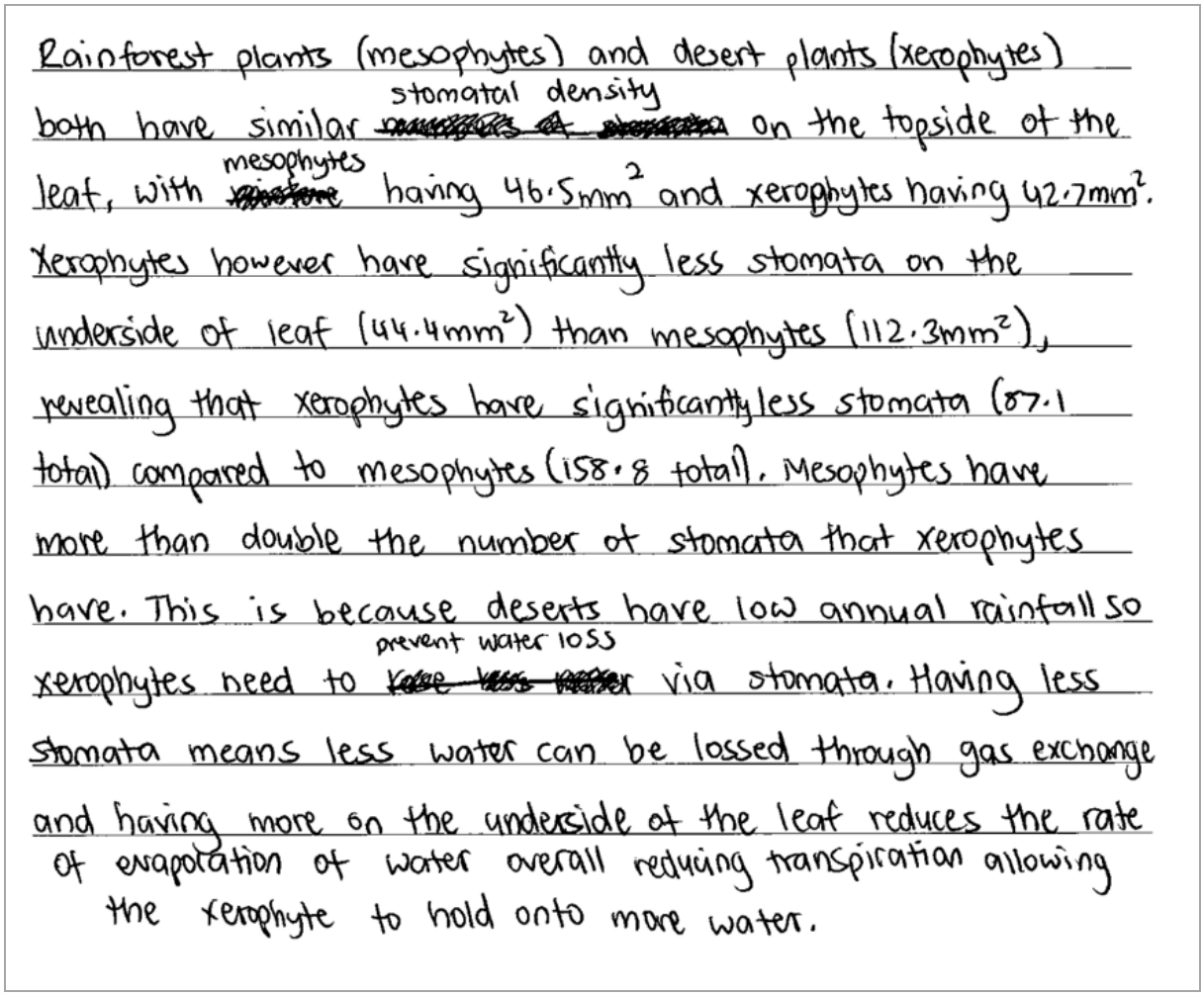
The following excerpt is from Question 26 in AS Paper 1. It required students to compare the stomatal density of rainforest plants and desert plants.

Effective student responses identified:

- a similarity
- a difference
- the significance.

This excerpt has been included:

- to demonstrate identification of the significance as a feature of *compare* questions.



Rainforest plants (mesophytes) and desert plants (xerophytes) both have similar stomatal density on the topside of the leaf, with mesophytes having 46.5mm^2 and xerophytes having 42.7mm^2. Xerophytes however have significantly less stomata on the underside of leaf (44.4mm^2) than mesophytes (112.3mm^2), revealing that xerophytes have significantly less stomata (87.1 total) compared to mesophytes (158.8 total). Mesophytes have more than double the number of stomata that xerophytes have. This is because deserts have low annual rainfall so xerophytes need to prevent water loss via stomata. Having less stomata means less water can be lossed through gas exchange and having more on the underside of the leaf reduces the rate of evapotation of water overall reducing transpiration allowing the xerophyte to hold onto more water.

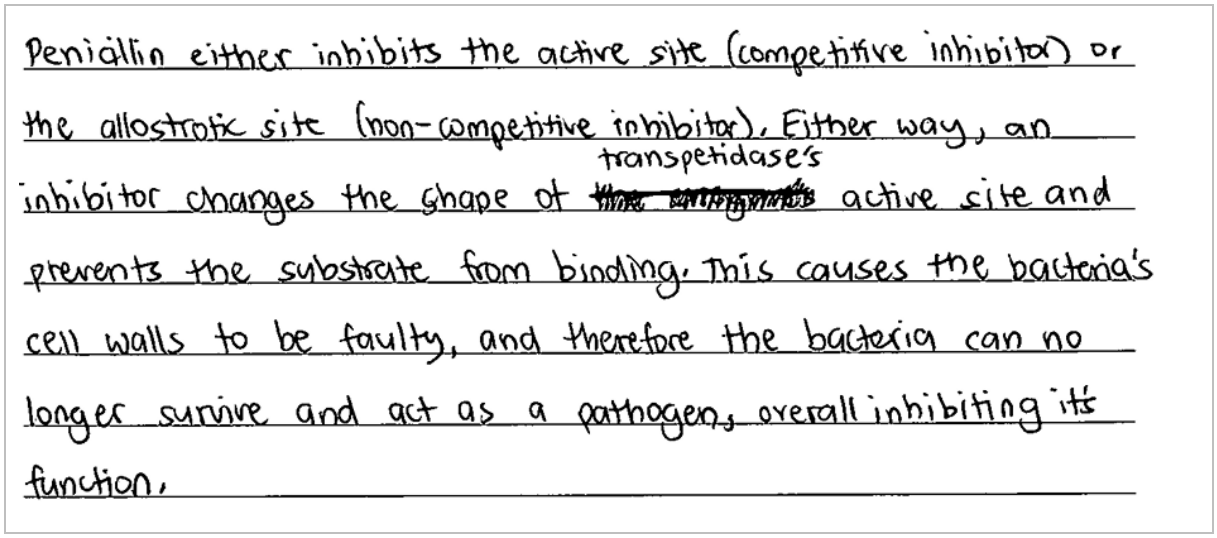
The following excerpt is from Question 4c) in AS Paper 2. It required students to explain how penicillin affects the structure and action of transpeptidase to form bacterial cell walls.

Effective student responses explained:

- how penicillin affects the shape of the active site
- the effect on the action of transpeptidase

This excerpt has been included:

- to demonstrate the *explain* cognition.


Penicillin either inhibits the active site (competitive inhibitor) or the allostroic site (non-competitive inhibitor). Either way, an inhibitor changes the shape of transpetidase's active site and prevents the substrate from binding. This causes the bacteria's cell walls to be faulty, and therefore the bacteria can no longer survive and act as a pathogen, overall inhibiting its function.

Practices to strengthen

When preparing students for external assessment, it is recommended that teachers:

- provide teaching and learning opportunities to respond appropriately to questions by identifying key elements that must be included in the response, e.g. identifying the principles of natural selection when asked to use them in explanations of similarities between two species
- emphasise the importance of cognitive verbs and number of marks when planning responses, e.g. a comparison requires identification of similarities, differences and their significance
- explore opportunities for students to respond to stimulus appropriately rather than making generalised statements
- consider connections between syllabus subject matter to integrate deepened understanding e.g. how coevolution of two species may occur.