

Psychology subject report

2021 cohort

February 2022

ISBN

Electronic version: 978-1-7438-190-6



© State of Queensland (QCAA) 2021

Licence: <https://creativecommons.org/licenses/by/4.0> | **Copyright notice:** www.qcaa.qld.edu.au/copyright — lists the full terms and conditions, which specify certain exceptions to the licence. |

Attribution: '© State of Queensland (QCAA) 2021' — please include the link to our copyright notice.

Other copyright material in this publication is listed below.

1. Student responses in this report are excluded from the CC BY 4.0 licence.

Queensland Curriculum & Assessment Authority
PO Box 307 Spring Hill QLD 4004 Australia
154 Melbourne Street, South Brisbane

Phone: (07) 3864 0299

Email: office@qcaa.qld.edu.au

Website: www.qcaa.qld.edu.au

Contents

Introduction	1
Audience and use	1
Report preparation.....	1
Subject data summary	2
Subject completion	2
Units 1 and 2 results.....	2
Units 3 and 4 internal assessment (IA) results	2
Total marks for IA	2
IA1 marks.....	3
IA2 marks.....	4
IA3 marks.....	5
External assessment (EA) marks	6
Final subject results	7
Final marks for IA and EA.....	7
Grade boundaries.....	7
Distribution of standards.....	7
Internal assessment	8
Endorsement	8
Confirmation	8
Internal assessment 1 (IA1)	10
Data test (10%).....	10
Assessment design	10
Assessment decisions	12
Internal assessment 2 (IA2)	13
Student experiment (20%)	13
Assessment design	13
Assessment decisions	15
Internal assessment 3 (IA3)	21
Research investigation (20%).....	21
Assessment design	21
Assessment decisions	23
External assessment	29
Summative external assessment (EA) — Examination (50%).....	29
Assessment design	29
Assessment decisions	30

Introduction

Despite the challenges brought about by the COVID-19 pandemic, Queensland's education community can look back on 2021 with satisfaction at having implemented the first full assessment cycle in the new Queensland Certificate of Education (QCE) system. That meant delivering three internal assessments and one external assessment in each General subject.

This report analyses that cycle — from endorsing summative internal assessment instruments to confirming internal assessment marks, and designing and marking external assessment. It also gives readers information about:

- applying syllabus objectives in the design and marking of internal and 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 of best practice where relevant, possible and appropriate.

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 external assessment.

The report is publicly available to promote transparency and accountability. Students, parents, community members and other education stakeholders can learn about the assessment practices and outcomes for General subjects (including alternative sequences (AS) and Senior External Examination (SEE) subjects, where relevant) and General (Extension) 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 data summary

Subject completion

The following data includes students who completed the General subject or AS.

For the purposes of this report, while the 2021 summative units for the AS are AS units 1 and 2, this information will be included with the General summative Units 3 and 4.

Note: All data is correct as at 17 December 2021. Where percentages are provided, these are rounded to two decimal places and, therefore, may not add up to 100%.

Number of schools that offered the subject: 142.

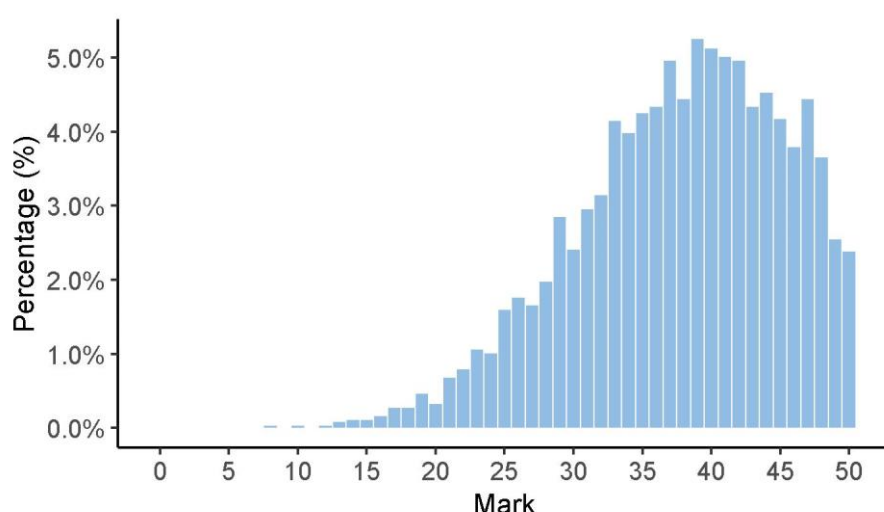
Completion of units	Unit 1	Unit 2	Units 3 and 4
Number of students completed	4432	4163	3642

Units 1 and 2 results

Number of students	Satisfactory	Unsatisfactory
Unit 1	4105	327
Unit 2	3833	330

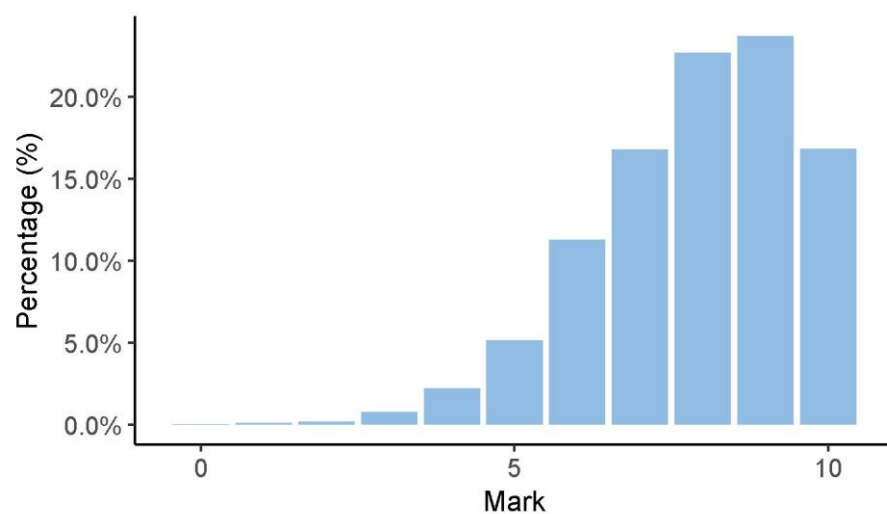
Units 3 and 4 internal assessment (IA) results

Total marks for IA

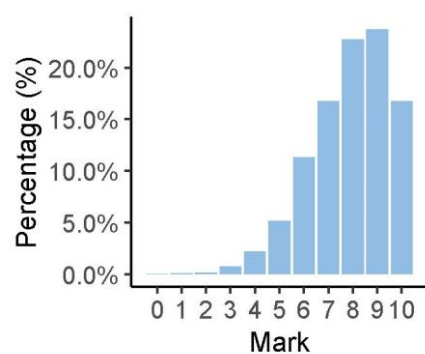


IA1 marks

IA1 total

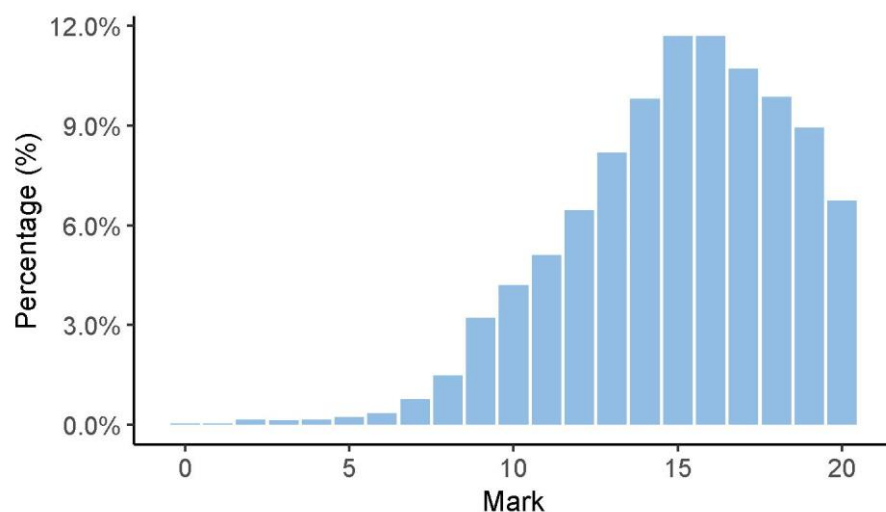


IA1 Criterion: Data test

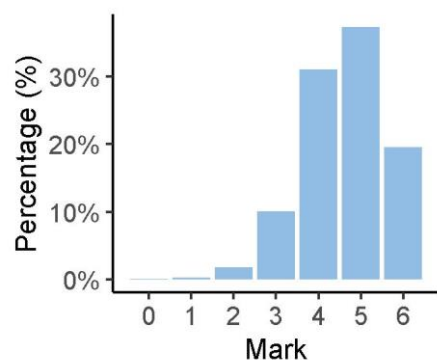


IA2 marks

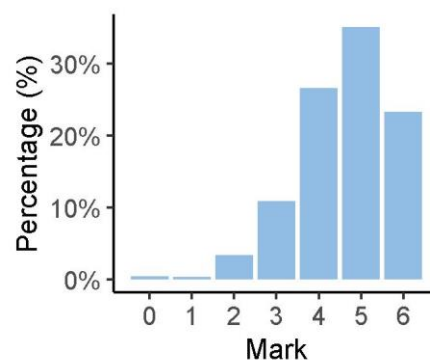
IA2 total



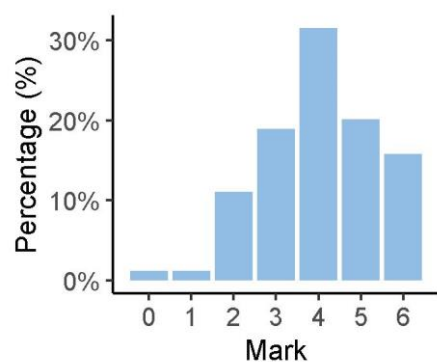
IA2 Criterion: Research and planning



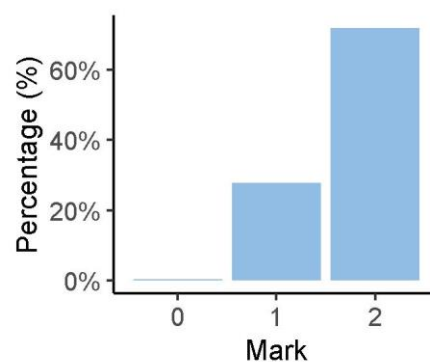
IA2 Criterion: Analysis of evidence



IA2 Criterion: Interpretation and evaluation

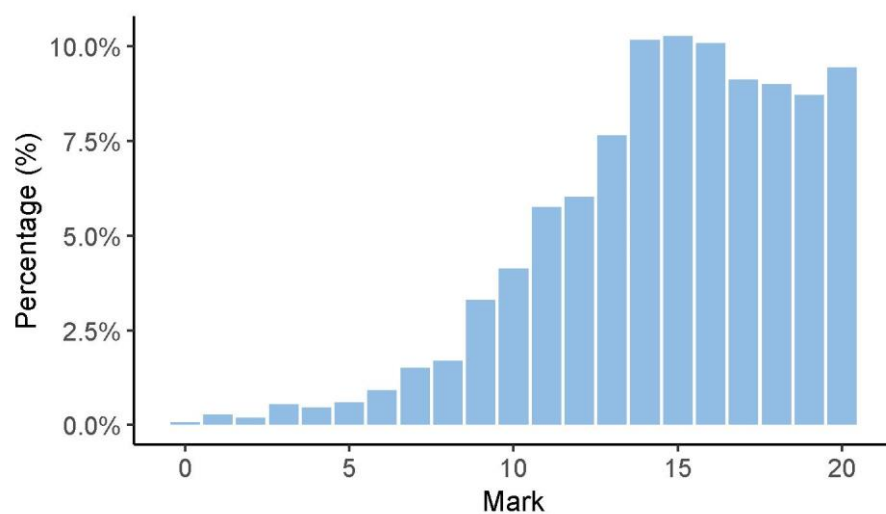


IA2 Criterion: Communication

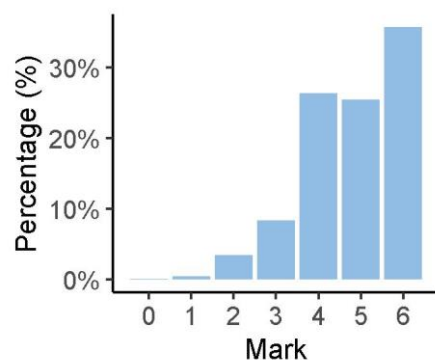


IA3 marks

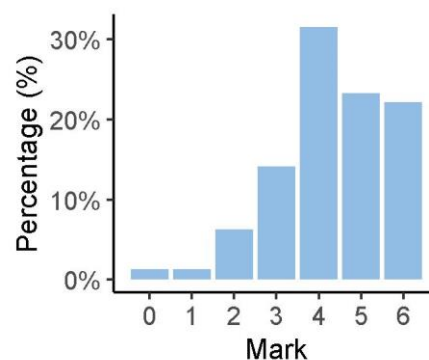
IA3 total



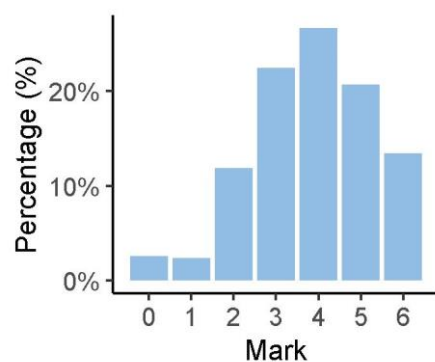
IA3 Criterion: Research and planning



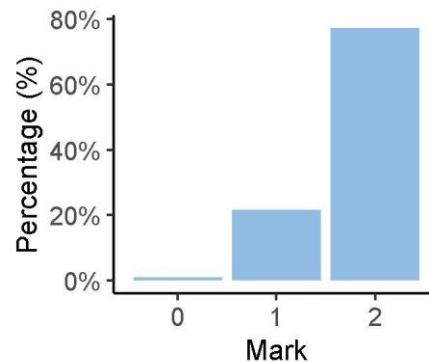
IA3 Criterion: Analysis and interpretation



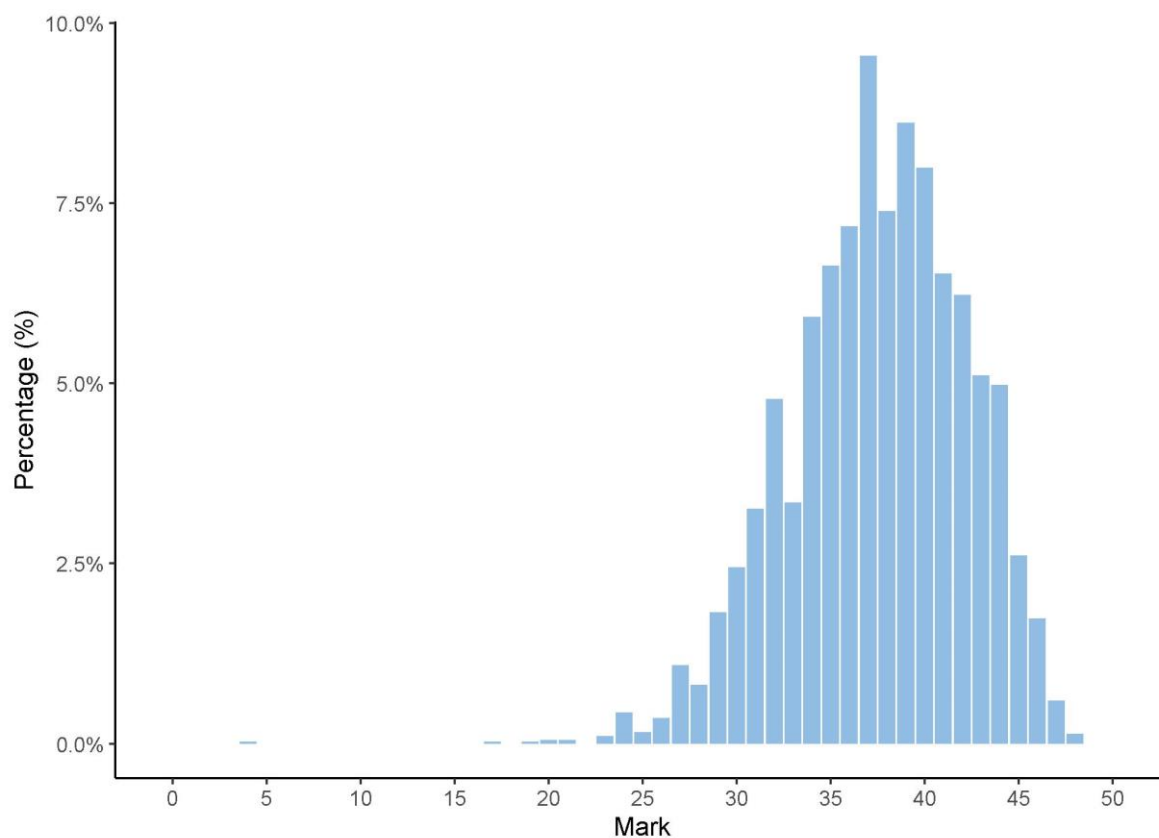
IA3 Criterion: Conclusion and evaluation



IA3 Criterion: Communication

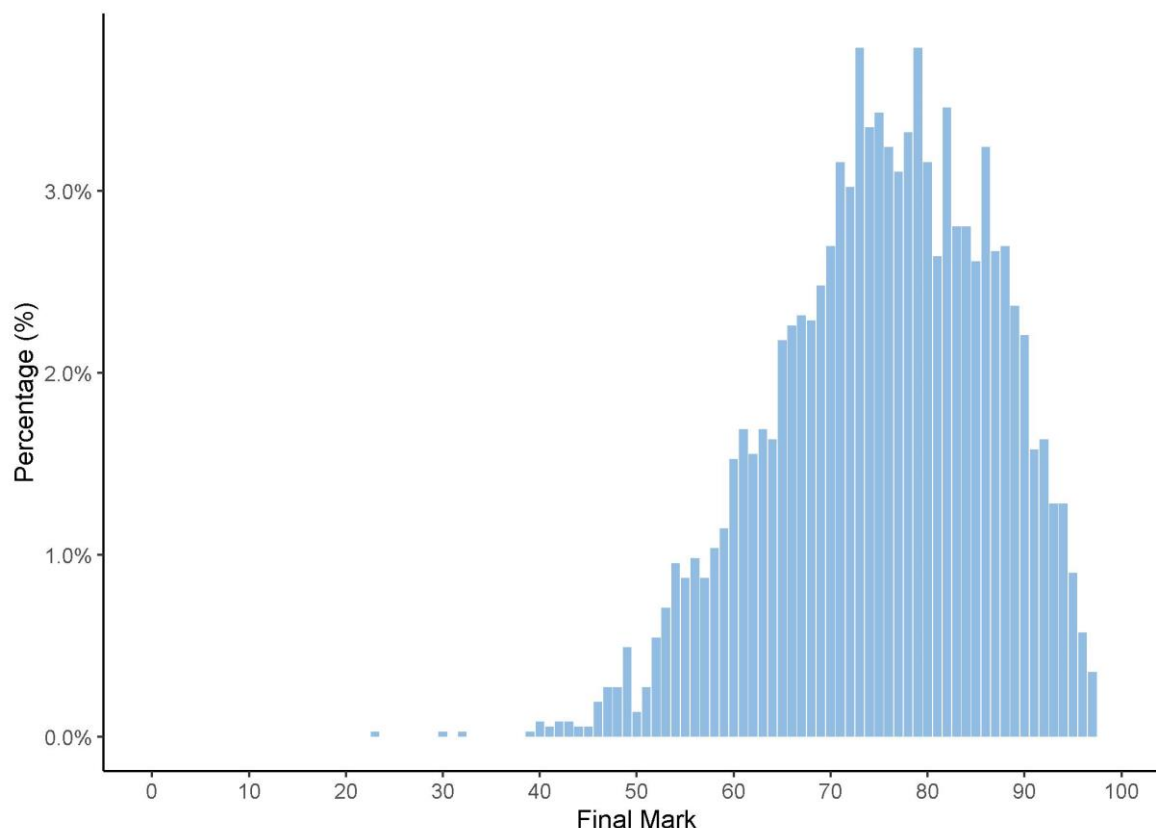


External assessment (EA) marks



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	100–85	84–70	69–48	47–22	21–0

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	850	1741	1017	36	0



Internal assessment

The following information and advice pertain 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 the quality assurance tools for detailed information about the assessment practices for each assessment instrument.

Percentage of instruments endorsed in Application 1

Number of instruments submitted	IA1	IA2	IA3
Total number of instruments	142	142	141
Percentage endorsed in Application 1	15%	73%	74%

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 ISMG and are used to make decisions about the cohort's results. If further information is required about the school's application of the ISMG to finalise a confirmation decision, the QCAA requests additional samples.

Schools may request a review where an individual student's confirmed result is different from the school's provisional mark in one or more criteria and the school considers this result to be an anomaly or exception.

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	141	827	0	100%
2	141	879	240	82.98%
3	140	864	202	83.57%



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	117
Authentication	0
Authenticity	6
Item construction	16
Scope and scale	39

*Each priority might contain up to four assessment practices.

Total number of submissions: 142.

Effective practices

Validity priorities were effectively demonstrated in assessment instruments that:

- featured a variety of datasets clearly derived from the subject matter of the relevant unit, e.g. mandatory and suggested practicals and related subject matter
- featured items with only one cognition for the objective being assessed
- featured a variety of cognitions throughout all datasets to enable students to demonstrate all aspects of the syllabus objectives
- include a marking scheme that clearly and consistently matches each mark to an important feature in the expected response, e.g. one mark for working and one mark for calculating the correct value.

Practices to strengthen

It is recommended that assessment instruments:

- contain items that are clearly aligned with the corresponding objective by using an appropriate cognitive verb and requiring an appropriate type of response, e.g. in an Objective 3 item, identify a trend in the dataset. Teachers should refer to the Mark allocations table in Syllabus section 4.5.1 for guidance on the appropriate cognitive verbs and responses associated with each objective
- include items that are closely aligned to the key research skills in Psychology in Syllabus section 1.2.5
- use unseen datasets derived from practicals performed in class and that are appropriately different from data sets available in QCAA samples and known publishers in the public domain
- do not use items that assess Objective 5 — investigate phenomena, e.g. items relating to hypothesis construction and limitations of experimental design, which are outside the scope of the data test.

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	28
Language	30
Layout	28
Transparency	17

*Each priority might contain up to four assessment practices.

Total number of submissions: 142.

Effective practices

Accessibility priorities were effectively demonstrated in assessment instruments that:

- featured a consistent layout, e.g. labelled axes on graphs and titles on tables
- featured clearly legible datasets, e.g. data points on correlational graphs and columns on bar graphs that were identifiable
- featured a clearly identifiable cognition in the instruction
- featured language that was free from cultural, gender or socio-economic bias
- featured items that succinctly provided the aim and methodology of the study before presentation of the data
- featured clear communication that directed students to the relevant data or table in order to respond to the question.

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	Data test	100%	0%	0%	0%

Effective practices

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

- school-developed marking schemes clearly matched each mark to a valued feature of the expected response and clearly explained how marks (including part marks) were awarded
- schools applied their marking schemes consistently across cohorts.

Samples of effective practices

There are no student response excerpts because either the student/s did not provide permission or there were third-party copyright issues in the response/s.

Practices to strengthen

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

- if necessary, schools update marking schemes to identify how alternative responses were marked
- schools check that mark totals and percentages have been determined correctly
- schools use the percentage cut-offs from the ISMG to determine the final mark out of 10
- schools implement internal quality assurance processes (e.g. cross marking) to ensure intra-marker and inter-marker reliability.

Additional advice

- Schools should label the updated marking scheme for ease of access during the confirmation process.



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	14
Authentication	10
Authenticity	0
Item construction	10
Scope and scale	3

*Each priority might contain up to four assessment practices.

Total number of submissions: 142.

Effective practices

Validity priorities were effectively demonstrated in assessment instruments that:

- featured clear authentication strategies e.g. a declaration of authenticity
- featured mandatory or suggested practicals from Unit 3 as experiments for students to use as the basis for their methodology and research question
- featured check points to monitor student progress throughout the task, e.g. specific weeks to select modifications, complete risk assessment, collect and analyse data, submit draft, submit final response
- featured scaffolding that enabled students to address the components of the task without directing them to a pre-determined response
- featured a clear example of developing a research question that related to Unit 3 subject matter.

It is recommended that assessment instruments:

- direct students to practicals that have been modelled in class, e.g. mandatory or suggested practicals that provide relevant data to process, such as an experimental research design to investigate the effect of learning environment on memory, replicating aspects of the 1998 investigation by Grant et al. in Syllabus section 4.5, Topic 3: Memory
- avoid the use of practicals that cannot be easily replicated in a classroom setting, e.g. Bugelski and Alampay (1961) and Hudson (1960) in Syllabus section 4.4, Topic 2: Visual perception
- include scaffolding that reflects the requirements of the syllabus and supports the completion of the task without leading students to a predetermined response
- include an indication of how students can work collaboratively and how this will be managed by indicating which sections of the assessment can be completed individually and which sections are to be completed as a group
- ensure that the mandatory and suggested practicals align to the topics stated in the conditions section of the task.

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	7
Layout	2
Transparency	5

*Each priority might contain up to four assessment practices.

Total number of submissions: 142.

Effective practices

Accessibility priorities were effectively demonstrated in assessment instruments that:

- featured clear communication of task elements using succinct and accessible language
- featured clear, appropriate headings with only relevant information in each section to provide guidance for students
- featured check points that were specific and provided clear directions for students.

Practices to strengthen

It is recommended that assessment instruments:

- are proofread for spelling, grammar and typographical errors
- avoid repetition of elements of the task in different sections.

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	89.36%	10.64%	0.00%	0.00%
2	Analysis and interpretation	91.49%	8.51%	0.00%	0.00%
3	Interpretation and evaluation	87.23%	12.77%	0.00%	0.00%
4	Communication	97.87%	1.42%	0.71%	0.00%

Effective practices

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


- in the Research and planning criterion
 - a considered rationale developed the research question from the relevant psychological concepts from Unit 3 subject matter, clearly identifying the experiment to be modified, communicating the experiment's purpose and introducing research to support altered variables
 - a hypothesis was accompanied by a specific and relevant research question, including operationalisation of variables and target population, which was answerable within the required response length
 - justified modifications to the methodology, explicitly identifying meaningful refinements, redirections or extensions of a practical, ensuring collection of sufficient data to draw valid conclusions
 - considered management of risks and ethical or environmental issues was explicitly identified within the context of the experiment
- in the Interpretation and evaluation criterion
 - justified discussion of the reliability and validity of the process by
 - providing an explanation of research informed by types of reliability and validity rather than a simple list of issues
 - referring to the uncertainty and limitations identified in the analysis of evidence
 - suggested improvements to the experiment were logically derived from the uncertainty and limitations of the evidence identified in the analysis, to improve reliability and validity of the experiment
 - suggested extensions to the experiment were purposefully explained to identify further scope for exploration or redirection of the research question.

Samples of effective practices

The following are excerpts from a response that illustrate the characteristics for the criterion at the performance level indicated. The excerpts may provide evidence of more than one criterion. The characteristics identified may not be the only time the characteristics have occurred throughout a response.

These student response excerpts have been included:

- to demonstrate the following characteristics in the Research and planning criterion
 - a considered rationale that clearly identifies the original practical, the purpose of the experiment and an explanation of the redirected variables that is informed by research
 - a *specific and relevant* research question that includes operationalisation of variables and the target population
 - modifications to the methodology that are justified by explicitly identifying meaningful refinements and redirections, explaining the purpose of modifications and explicitly operationalising variables.

<p>Research and planning (5–6 marks)</p> <ul style="list-style-type: none"> • a considered rationale for the experiment 	<div style="display: flex; align-items: flex-start;"> <div style="margin-right: 10px;">  <p>Rationale</p> </div> <div> <p>Memory is the ability to acquire, store, retain and retrieve information (Cherry, 2020). It can be affected by a number of factors, including motivational and emotional factors, the mood or state of the person and their physical circumstance or context (Goddard, 2012). Context-dependent memory refers to the improved ability to retrieve memories when a person's context or circumstances match for both the original encoding of the memory and the retrieval of it (Walinga & Stangor, 2010). For example, it is usually much easier for people to recall childhood memories when in the place they grew up.</p> <p>Grant et al. performed an experiment investigating context-dependent memory, in which participants aged 17-56 studied in either a quiet or noisy environment and were then tested in one of those two environments (Grant, et al., 1998). The test contained multiple choice and short-answer questions, testing two types of the participants' memory: recognition memory, a subset of declarative memory which involves the ability to identify correct responses from a set of alternatives based on previous experience, and recall memory, the retrieval of stored information using minimal cues (Burton, et al., 2019). Overall, the participants who had matching study and testing environments outperformed those whose environments did not match on both types of questions, indicating a significant effect of context on recognition and recall memory. Another context-dependent memory study was performed by D.R. Godden and A.D. Baddeley which produced similar results (Godden & Baddeley, 1975). 18 adult divers learned a list of words on dry land and underwater and recalled the words in either their original learning environment or the alternate environment. Those who had the same learning and testing environments again significantly outperformed those with unmatched environments.</p> <p>Several studies have investigated context-dependent memory and its effect on the performance of participants in a testing environment, but few have tested the age group of 16 to 18-year-olds, instead focusing solely on adults. As people between the ages of 16 and 18 are typically senior high school students, studying and exams are heavily prevalent in their lives (Morin, 2021). Some students prefer to study on paper, reading physical textbooks and taking handwritten notes, while others read online and type their notes (Bothwell, 2017). Seeing as most, if not all senior exams are taken on paper (Lee, 2015), the students' study methods and whether they match their paper exams may have an effect on their performance due to the effect of context-dependent memory.</p> <p>To further investigate context-dependent memory on this particular age group, Grant et al's experiment will be modified to target senior students in Years 11 and 12. Some students will study on paper and the others online, yet all will take a multiple-choice test on paper, testing their recognition memory. Their study and testing environment will be a high school science laboratory.</p> </div> </div>
---	---

**Research and planning
(5–6 marks)**

- a specific and relevant research question

**Research and planning
(5–6 marks)**

- justified modifications to the methodology

Research question

Will senior students in Queensland who both study and take a multiple choice (memory recognition) test on paper outperform students who study online and take the same test on paper?

Modifications to methodology:

Grant et al.'s experiment was modified in the following ways.

Variables

- The independent variable of studying with or without noise was altered to studying on paper or online and testing on paper, but the focus on performance of students with matched vs unmatched conditions remained the same.
- While the original experiment had 2 independent variables – studying in quiet or in noise, and testing in quiet or in noise, the modified experiment will only have one – studying on paper or online, as all tests will be taken on paper. This will both simplify the experiment and reflect the typical testing environment of high school students as senior exams are rarely online whereas the students' study environments are more likely to vary.
- There will only be one dependent variable rather than two – the participants' performance on a multiple-choice test – testing only their recognition memory to simplify and narrow the focus of the investigation.
- Study time will be controlled, with each student receiving the same amount of time to read the given material to remove the effect that the variable of time reading may have on the results.

Participants

- The population was changed to senior high school students in Queensland as there has been little to no research investigating context-dependent memory on this age group.
- The sample size will be 28 students, 14 for each condition, due to the number of students available for testing.

This student response excerpt has been included:

- to show a justified discussion of reliability and validity of the experimental process that links limitations to the types of reliability and validity affected
- to demonstrate suggested improvements to the experiment that are logically derived from the limitations of the evidence identified in the analysis to improve the reliability and validity of the experiment. Suggested extensions to the experiment are purposefully explained to identify further scope for exploration or redirection of the research question.

**Interpretation and evaluation
(5–6 marks)**

- justified discussion of the reliability and validity of the experimental process

LIMITATIONS WITHIN THE EXPERIMENTAL PROCESS

A lack of reliability and validity in the experimental process is evidenced through a small sample size, large standard deviation and error, ^{and} the lack of population, ecological, content, construct, external and internal validity were observed from the evaluation and analysis.

Large standard deviations ($NCR = 2.85$, $CR = 3.23$) were observed for each condition. A large standard deviation indicates that data points are widely dispersed around the mean, suggesting that extraneous variables are not fully controlled thus decreasing accuracy. High discrepancies in the sample set decreases the internal reliability and, this high standard deviation increases the uncertainty in the data.

Data that is widely dispersed within its statistical boundaries and wide parameters (population) can reveal an unreliable experimental methodology. Although experimenters sought to control most extraneous variables through a planned methodology, issues arose due to the emergence of uncontrolled factors. The uncertainty in the data is probably attributed to natural participant variability. When exposed to a word list at a point in time individual's proficiency varies, magnified effect in a small sample size. To attempt to control this, experimenters used convenience sampling and matched participant's design. A small sample size ($N = 38$) and an independent group design most likely caused the dispersion (lack of internal reliability) in the data. *why?*

The refinements made to the original experiment's methodology, the randomisation of the rhyming words to minimise bias, meant that the most suitable rhyming word was not always generated. Thus, the lack of a standardised rhyming list may have affected test-retest reliability. Trends in recall may have been affected by personal affiliation to the words or varying difficulty of the words to be remembered, causing outliers, increasing bias so ✓

decreasing internal validity. Moreover, the ability to accurately apply these results in other real-world contexts is limited, decreasing ecological validity.

The population validity is low due to the small sample size ($N = 38$), and unrepresentative sample (only Grade 11 students aged between 15 and 16 years). This small sample size and unrepresentative sample do not represent the general population and so decreases the ecological validity of the experiment.

Interpretation and evaluation (5–6 marks)

- suggested improvements and extensions to the experiment that are logically derived from the analysis of evidence

SUGGESTED IMPROVEMENTS AND EXTENSIONS FOR THE EXPERIMENTAL PROCESS

By analysing the evidence obtained, it was clear that the experimental processes lacked reliability and validity and the following improvements and extensions are suggested to improve reliability and validity of the experimental process.

Despite the significant result, to reduce the possibility of Type 1 errors and lack of internal validity, the following improvements should be made; use of distraction task for participants (i.e counting backward from 100 by sevens) to eliminate rehearsal of the words. Moreover, the use of other cues (i.e colour or categories) tested separately in a laboratory environment - to eliminate distractions and copying – ensure the difference in recall is due solely to the cues; rather than the type of cue, ability to rehearse the words or ability to cheat.

To improve the internal reliability of the sample, the sample size should be increased to > 100 to ensure that each individual participant contributes less to the overall result, decreasing the dispersion.

A further improvement would be for participants to complete a pre-test for their memory ability, paired, and then allocated to either the CR or NCR conditions. This measure further decreases the chance of natural participant variability affecting the results, making it a more construct and concurrently valid experiment to test memory recall.

An extension to increase the ecological and population validity would be to test a more diverse population group through random sampling techniques (i.e randomised sample of children, adolescents, elderly). This would allow experimenters to discover if the effect of cues on recall was age or life-stage dependent. To improve test-retest reliability – standardised word lists and rhyming recall words (considered reliable by other psychological researchers and construct validity).

Practices to strengthen

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

- in the Analysis of evidence criterion
 - correct and relevant processing of data is determined by the appropriate use of measures to calculate statistical difference and/or significance such as confidence intervals and inferential statistics rather than inappropriate measures such as standard deviation and standard error
 - thorough identification of trends, patterns or relationships involves the following analysis techniques (where relevant)
 - explanation, rather than re-statement, of descriptive statistics, e.g. mean, median, standard deviation and standard error
 - identification of trends within correlation (if appropriate to answer the research question), including the strength and direction of the relationship based on Pearson's correlation coefficient or Spearman's rank correlation coefficient
 - measures to identify statistical difference and/or significance, including confidence intervals and appropriate inferential statistics for the sample and conditions, e.g. Mann-Whitney U test, T-test, Wilcoxon signed-ranked test with explanation of selection for appropriateness

- thorough and appropriate identification of uncertainty and limitations of evidence involves the following descriptions (where relevant)
 - uncertainty within the evidence is identified based on error bars, type I & II errors, management of outliers within the data collected and instrumental data measurement uncertainties
 - limitations within evidence are based on sample size, population and sample used and implementation of methodology.

Additional advice

- Schools must use the ISMG from the syllabus without making any changes to wording or formatting.
- According to the best-fit process for using the ISMG, when an equal number of characteristics are identified across two performance levels (e.g. in the Analysis of evidence criterion), the mark should reflect that balance of evidence in the response. Teachers should refer to the *Making judgements webinar* resource available on the QCAA Portal for additional guidance.
- Through the mandatory and suggested practicals, students should have an opportunity to practise
 - relevant data processing techniques that can be used to identify trends, patterns, relationships and the uncertainty or limitations of data
 - interpreting statistical significance using appropriate measures (e.g. confidence levels) rather than inappropriate measures (e.g. standard deviation).
- Strategies outlined in the QCE and QCIA policy and procedures handbook are administered to
 - manage response length to ensure that student responses meet the conditions of the syllabus
 - promote academic integrity to ensure that student responses clearly demonstrate students' own achievements.



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	18
Authentication	2
Authenticity	0
Item construction	9
Scope and scale	16

*Each priority might contain up to four assessment practices.

Total number of submissions: 141.

Effective practices

Validity priorities were effectively demonstrated in assessment instruments that:

- used claims that were simple and direct statements that enabled the development of a research question aligned to Unit 4 subject matter, e.g. 'the media can decrease aggression'
- used claims that could generate multiple unique student responses, e.g. 'cultural norms influence social identity'
- used scaffolding that directed students to address all elements of the task.

Practices to strengthen

It is recommended that assessment instruments:

- use claims that only have a single assertion, e.g. 'diffusion of responsibility occurs in all social situations' rather than 'diffusion of responsibility explains the reluctance of people to go to the aid of strangers who are bullied'
- consider the school context and resources available to students to support them in finding evidence for their research question
- include the task requirements in the task section, not the scaffolding 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	4
Layout	1
Transparency	0

*Each priority might contain up to four assessment practices.

Total number of submissions: 141.

Effective practices

Accessibility priorities were effectively demonstrated in assessment instruments that:

- used claims written with accurate spelling and grammar
- provided clear instructions that aligned with syllabus objectives and the ISMG
- had checkpoints that were sufficiently spaced and had clear completion expectations to support students in the successful construction of the task, e.g. specific weeks to develop research question, collect and analyse data, submit draft, submit final response

Practices to strengthen

It is recommended that assessment instruments:

- avoid using colloquial language or popular catch phrases, e.g. 'sticks and stones may break my bones but words will never hurt me'
- consider the cultural appropriateness of the claim within a broader societal context
- avoid using claims that are moral assertions or have a moral bias, e.g. 'ethical principles in psychology have hindered psychological research'.

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	94.29%	3.57%	1.43%	0.71%
2	Analysis and interpretation	93.57%	5.71%	0.71%	0.00%
3	Conclusion and evaluation	85.71%	12.14%	2.14%	0.00%
4	Communication	97.86%	0.71%	1.43%	0.00%

Effective practices

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

- in the Research and planning criterion
 - a considered rationale clearly connected the research question to Unit 4 subject matter
 - a specific research question was explicit enough to be answered within the required response length
 - sufficient and relevant sources were drawn from a variety of scientifically credible sources as indicated by the syllabus, e.g. books by well credentialled scientists, websites of independent research bodies, etc.
 - judgements about the sufficiency of sources were made by referring to the sources used throughout the response, not just in the rationale
- in the Conclusion and evaluation criterion
 - a justified conclusion directly addressed the research question and was supported by evidence from credible sources
 - discussion of the quality of evidence was related to the research question
 - suggested improvements and extensions were related to the claim
- in the Communication criterion
 - appropriate use of referencing conventions ensured that a consistent style of in-text referencing was used throughout the response.

Samples of effective practices

The following is an excerpt from a response that illustrates the characteristics for the criterion at the performance level indicated. The excerpt may provide evidence of more than one criterion.

The characteristics identified may not be the only time the characteristics have occurred throughout a response.

This student response excerpt has been included:

- as an example, in the Research and planning criterion, of a rationale that is considered and a research question that is clearly developed from the claim
- as an example, in the Analysis and interpretation criterion, of sufficient and relevant evidence clearly addressing the research question
- as an example, in the Conclusion and evaluation criterion, of justified conclusions using data to answer the research question and credible findings being extrapolated to address the claim.

**Research and planning
(5–6 marks)**

- a considered rationale identifying clear development of the research question from the claim
- a specific and relevant research question

✓ **Claim:** Is Prosocial or Altruistic behaviour innate?

Rationale:

✓ The concept of altruism was coined as the antithesis to egoism by Auguste Comte in the 19th century
 ✓ (Britannica, 2019). Since then, the definition of altruism has continued to evolve. As research has
 begun to find relationships between genetics and certain behaviours, psychologists have continued to
 ✓ try to find out whether the virtue is innate, or because of social influence. However, this research has
 suffered due to conceptual confusion, resulting in failed attempts at answering whether altruism is
 ✓ innate (Batson & Shaw, 1991).

To answer the claim "Is prosocial or altruistic behaviour innate?" clear definitions need to be provided
 to cover the shortcomings of prior research. For the purposes of this investigation altruism will be
 defined as "a motivational state with the ultimate goal of increasing another's welfare" (Batson &
 ✓✓ Shaw, 1991). Altruistic behaviours can include doing something to help another person with no
 expectation of reward, forgoing things that may bring personal benefits if they create costs for others,
 sharing resources (potentially when scarce) and helping someone despite potential personal risks
 ✓ (Cherry, 2018). Extreme forms of altruism such as self-sacrificial behaviour in dangerous situations is
 impossible to observe due to ethical restraints. This leads to much smaller examples of altruistic acts
 ✓ in children being extrapolated to determine the innateness of altruism.

Because of observer bias in laboratory settings and the exposure to societal influence, studies on
 altruism have focused on the behaviours of young children (usually between 1-2 years old) to truly
 ✓ find if this behaviour is innate (Barragan & Dweck, 2014) as its generally accepted that this age is less
 ✓ effected by societal influence. However, the innate factor may be affected due to the presence of
 reciprocal activities present in experiments with young children, as former research found that without
 ✓ reciprocal activities it found little subsequent altruism. This calls into question the claim that human
 children are altruistic by nature, and the belief that socialization cannot account for altruistic acts
 (Barragan & Dweck, 2014).

citation needed

**Analysis and interpretation
(5–6 marks)**

- the identification of sufficient and relevant evidence

Research Question:

As such, the proposed research question is as follows:

Do children between 1 to 2 years of age exhibit the altruistic behaviours of helping and sharing with others without the presence of rewards or reciprocal activities? ✓

Finally, a study conducted by Barragan and Dweck (2014) found that simple acts of reciprocal play found higher rates of subsequent altruistic behaviours of helping in 1–2-year-old children, whereas friendly but non-reciprocal play didn't. ✓

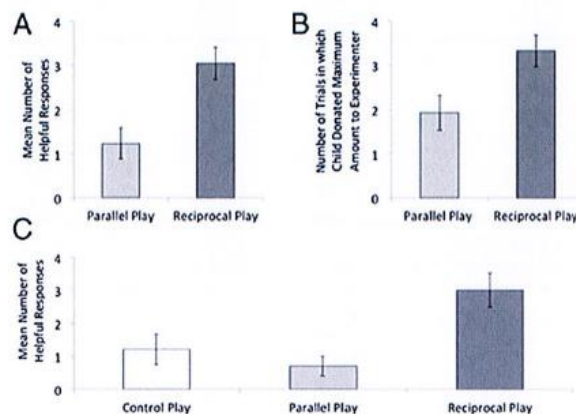


Figure 4. Effect of reciprocal play manipulation on helping in study 1 (A) and study 2 (C) and on generosity in study 3 (B). Error bars represent the SEM. ✓

Cortes Barragan, R & Dweck, CS 2014, "Rethinking natural altruism: Simple reciprocal interactions trigger children's benevolence: Fig. 1.," *Proceedings of the National Academy of Sciences*, vol. 111, no. 48, pp. 17071–17074, DOI: 10.1073/pnas.1419408111. Used with permission

Infants were selected for either the first group ($n=34$) or the second group ($n=30$).

Thirty-four infants were selected for the first group and thirty for the second group. The first group was subjected to a warm-up period that involved reciprocal play, wherein the experimenter would pull out one set of toys and actively play with the child, such as rolling a ball between the two of them and hitting notes on a small musical toy together. They also performed parallel play, in which the experimenter and child would play with the same toys from 6 feet away, while the experimenter used positive reinforcement and playful dialogue towards the child. After this the test phase begun, the experimenter facilitated scenarios in which they needed help to reach an object (a block, a bottle, a clothespin, and a pencil) that was out of reach. It was found that children who had received reciprocal play significantly helped the experimenter complete the task at $p=0.001$. The second study group underwent the exact same conditions during warm up, but when test phase begun the experimenter was replaced by an assistant, which yielded little subsequent helping behaviours as the bond during play had been severed, as represented in figure 4.

One limitation for this experiment is the social cues relating to play, especially for children of this age, similarly to the other research papers. Children who engage in reciprocal play with parents of sibling on a positive basis often would be more likely to relate play with an experimenter to positive experiences and members of family. This could lead to a higher rate of altruistic behaviour as an expectation to people within your family. It also suffers from the potential for past experiences to have an influence like the other studies.

Finding that altruistic behaviour is significantly affected by reciprocal activities around play discredits the idea that humans are born naturally inclined to be altruistic towards strangers at no benefit to themselves. However, the limitations above in addition to the contesting evidence make this hard to generalise to the wider population.

word choice
Good links to the claim

Conclusion and evaluation (5–6 marks)

- justified conclusion/s linked to the research question
- extrapolation of credible findings of the research to the claim

Conclusion:

✓ The research provided by Warneken and Tomasello (2006) and Barragan, Brooks, and Meltzoff (2020), contradicts the research by Barragan and Dweck (2014). Warneken and Tomasello (2006) and

Barragan, Brooks, and Meltzoff both found significant evidence for altruistic behaviour in infants

✓ without the presence of rewards and reciprocal activities, but due to the limitations within the

✓ experimental methodologies and societal influence and the evidence provided by Barragan and

✓ Dweck (2014) that suggests it is a learned trait that is encouraged by result of personal connections.

The findings from these researchers suggests that further studies are needed to draw a proper

✓ conclusion on the innateness of Altruism.

✓ To answer the claim "Is Prosocial or Altruistic behaviour innate?" it is suggested further research in

needed to make a definitive conclusion that can be applied to the broader population.

Evaluation:

✓ The Methodologies used by researchers struggle to eliminate the possibility that the actions taken by

✓ the infants in their experiments are not a result from a learned behaviour. Due to the cognitive ability

of toddlers its difficult to design a scenario that can lead to natural problem solving that hasn't been

✓ part of their everyday life to this point. Additionally, the experiments all feature relatively high levels

of altruistic acts in the control groups of their experiments, which may mean poor execution of their

✓ methodologies, which decreases ecological validity.

✓ Some extensions could include broadening the food sharing experiment to other objects and items of

value to a child, such as toys or other items, and focusing experiments on the interactions between

✓ children their own age, as this may eliminate family values that a child may attribute to situations,

especially when seeing adults in a place of authority.

✓ Generalising these results is ineffective for representing a wider population as the research is very

✓ specific over short term, laboratory experiments, which don't represent all the ideals of altruism.

✓ Therefore, caution should be taken in generalising these findings to a wider population.

Practices to strengthen

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

- in the Analysis and interpretation criterion
 - sufficient and relevant evidence is
 - directly related to the research question
 - based on subject matter from Unit 4 of the syllabus
 - drawn from more than one source
 - thorough and appropriate identification of limitations of evidence should identify issues such as
 - weak points of the data with respect to the research question

- methodological limitations of the sources
- whether justified scientific arguments are supported by data from the sources.

Additional advice

- Strategies identified in the *QCE and QCIA policy and procedures handbook* should be administered to
 - manage response length to ensure that student responses meet the conditions of the syllabus
 - promote academic integrity to ensure that student responses clearly demonstrate their students' own achievements.
- Schools must use the ISMG from the syllabus without making any changes to wording or formatting.
- Strategies identified in the *QCE and QCIA policy and procedures handbook* to manage response length should be administered to ensure that student responses meet the conditions of the syllabus.



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.

Summative external assessment (EA) — 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 (25 marks)
- Paper 1, Section 2 consisted of short-response questions (29 marks)
- Paper 2, Section 1 consisted of short-response questions (40 marks).

The examination assessed subject matter from Units 3 and 4. Questions were derived from the context of:

- Localisation of function of the brain
- Visual perception
- Memory
- Learning
- Social psychology
- Interpersonal processes
- Attitudes
- Cross-cultural psychology

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

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 (25 marks)
- Paper 1, Section 2 consisted of short-response questions (34 marks)
- Paper 2, Section 1 consisted of short-response questions (39 marks).

The AS examination assessed subject matter from AS Units 1 and 2. Questions were derived from the context of:

- The role of the brain

- Cognitive development
- Human consciousness and sleep
- Localisation of function of the brain
- Visual perception
- Memory
- Learning.

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

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.

General multiple choice item responses

There were 25 multiple choice items 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.

Question	A	B	C	D
1	3.55	87.98	6.01	2.21
2	5.62	1.54	90.16	2.49
3	0.48	3.77	91.45	4.14
4	6.21	6.68	12.8	74.11
5	17.03	69.11	11.66	1.96
6	70.51	6.23	12.24	10.85
7	3.52	2.94	88.68	4.64
8	62.85	6.23	20.94	9.73
9	11.32	40.84	23.15	24.29
10	1.62	13.08	5.54	79.45
11	4.05	47.44	21.08	27.17
12	79.09	11.15	2.82	6.71
13	3.02	8.97	4.5	83.37
14	70.76	2.12	4.56	22.34
15	65.5	1.59	1.98	30.75

Question	A	B	C	D
16	4.25	11.83	68.24	15.46
17	3.41	12.19	54.85	29.35
18	8.19	20.27	10.96	60.33
19	40.51	39.39	8.44	11.49
20	18.9	53.42	18.62	8.83
21	4.42	18.09	6.74	70.56
22	64.75	17.92	3.63	13.39
23	18.42	65.67	10.15	5.54
24	6.32	11.21	62.71	19.6
25	8.44	11.32	78.03	1.96

AS multiple choice item responses

There were 25 multiple choice items 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.

Question	A	B	C	D
1	91.3	4.35	4.35	0
2	8.7	8.7	43.48	39.13
3	4.35	73.91	17.39	4.35
4	21.74	8.7	39.13	30.43
5	17.39	4.35	73.91	4.35
6	4.35	69.57	8.7	17.39
7	0	0	91.3	8.7
8	56.52	4.35	30.43	8.7
9	4.35	30.43	39.13	26.09
10	0	17.39	4.35	78.26
11	39.13	43.48	8.7	8.7
12	4.35	4.35	4.35	86.96
13	17.39	39.13	8.7	34.78

Question	A	B	C	D
14	69.57	0	8.7	21.74
15	47.83	4.35	0	47.83
16	8.7	0	69.57	17.39
17	13.04	4.35	8.7	69.57
18	0	78.26	13.04	8.7
19	13.04	0	82.61	4.35
20	86.96	4.35	4.35	4.35
21	30.43	39.13	21.74	4.35
22	4.35	4.35	8.7	82.61
23	43.48	43.48	8.7	4.35
24	0	26.09	56.52	17.39
25	8.7	17.39	73.91	0

Effective practices

Overall, students responded well to:

- opportunities to apply their understanding of Unit 3 subject matter
- opportunities to apply their understanding of Unit 4 topic 4 subject matter
- short response questions with the cognitive verbs 'describe' and 'identify'.

The following excerpts have been selected to illustrate effective student responses in one or more of the syllabus assessment objectives. The characteristics identified may not be the only time the characteristics have occurred throughout a response.

Samples of effective practices

Short response

Assessment objective: 1

Paper 2

Question 1b

This question required students to describe extinction and spontaneous recovery and provide an example of each.

Effective student responses:

- described both extinction and spontaneous recovery
- identified examples of both extinction and spontaneous recovery from the experiment referenced in the syllabus and question.

This student response excerpt has been included:

- to demonstrate a high-level response that provides correct examples relevant to the syllabus.

Describe and explain
(4 marks)

Extinction ~~is the process~~ has occurred when an organism no longer provides the desired ~~an~~ response due to the lack of reinforcement after they display it. For example, if the birds were not ~~not~~ positively reinforced with seed many times after pecking the button, they no longer pecked it. Spontaneous recovery has occurred if an individual displays the desired behaviour even after a rest period without reinforcement. For example, even after ~~stop~~ gradually stopping pecking the button due to lack of reinforcement (seed), the birds ~~may spontaneously~~ sometimes would spontaneously exhibit the desired response (pecking).

Assessment objective: 3

Paper 1

Question 27

This question required students to distinguish between modelling and vicarious conditioning.

Effective student responses:

- identified a distinction between modelling and vicarious conditioning
- included a contrast word in the response.

This student response excerpt has been included:

- to demonstrate a high-level response that includes a relevant contrast word to show distinction.

Analyse evidence
(1 mark)

Modelling is the ^{of observation} process by which a role models behaviour ^{behaviour is learnt via the} being ~~to~~ observed whereas vicarious conditioning is the process by which behaviour is conditioned/learnt via observing the consequences/outcomes of role models behaviour.

Assessment objective: 4

Paper 1

Question 38

This question required students to identify an argument for and against elaborative rehearsal and support each argument with an example.

Effective student responses:

- provided arguments for and against elaborative rehearsal rather than referencing maintenance rehearsal
- provided examples that supported the arguments made.

This student response excerpt has been included:

- to demonstrate a high-level response that provides effective arguments for and against elaborative rehearsal as a strategy to improve memory, and examples to support the arguments made.

Interpret evidence
(4 marks)

QUESTION 38 (4 marks)

Provide an argument for and against using elaborative rehearsal as a strategy to improve memory. Support your response with an example for each argument.

~~But~~ An argument for using ^{elaborative} rehearsal to improve memory is that it allows for greater ^{of explicit information} encoding into long-term memory as it applies meaning to information so it can be better stored in semantic networks.

An example of when elaborative rehearsal is effective is in remembering facts such as ~~capitol~~ ^{capitol} cities by giving them meaning and relating them to an interesting story ~~also~~ or other information about the state they are the capitol of. ~~ET~~ An argument against the use of elaborative rehearsal is that it cannot be used when trying to remember procedural information or other implicit information. An example is trying to learn to touch-type, where remembering where all the letters are and becoming proficient at the action can only be achieved through repetition.

Practices to strengthen

It is recommended that when preparing students for external assessment, teachers consider:

- the skills associated with responding to cognitive verbs linked to subject matter, especially 'contrast', 'distinguish' and 'compare'
- the research referenced in the subject matter of the syllabus, rather than that referenced by outside sources, such as textbooks
- that the number of marks allocated per question reflects the number of statements, cognitions or calculations required to achieve full marks
- the multiple choice items where students answered incorrectly to ensure subject matter is sufficiently covered.