INTUITION EDUCATION

HSC 2026 · PREDICTED PRACTICE PAPER

2026 HSC Investigating Science

Answers & marking notes

Paper total: 100 marks

Sit the paper first. These are worked answers with the marking notes behind them — what earns the mark, and the traps the marking centre has flagged in past years. Where a note names a band, it describes what a top-band response does, not a guaranteed mark.

Published Aug 2026 · intuitioneducation.com.au/hsc-2026 · ai.intu.com.au

Answers & marking notes

Section I — answer key

Q	Answer	Note
1	C	The discussion evaluates the results, including reliability; results only presents them
2	A	Direct sensory/measured record; B is an inference, C a generalisation, D a prediction
3	B	van Helmont's willow experiment (his conclusion was reasonable but incomplete — CO_2 was unknown)
4	B	Observation → hypothesis → (accidental) culture → self-experimentation
5	B	The most serious stated risk is flammability; removing the ignition source addresses it — PPE options address lesser risks
6	B	Record one estimated digit beyond the smallest marked division; 36.50 implies unjustified precision
7	C	Salinity is deliberately changed; hatched count is the DV
8	C	A water-only control group isolates the vitamin as the cause — A improves reliability, B accuracy, D biases the data
9	B	Independent expert scrutiny of method and conclusions pre-publication
10	A	Fame in one field transfers unearned credibility to another — halo effect
11	B	Productivity responds to being observed, not to the changes themselves — Hawthorne effect
12	B	In-app, long-term-user, opt-in respondents are self-selected; dissatisfied users have already deleted the app
13	A	Technology (X-ray diffraction) supplied the evidence for the scientific model; B reverses it, C and D are science→technology
14	C	Wavefronts compressed on approach (higher f), stretched on recession (lower f) — Doppler effect
15	D	Boyle's law at constant T: $P_1V_1 = P_2V_2$; $100 \times 60 = P_2 \times 20 \Rightarrow 300$ kPa
16	C	Informed consent is an ethical safeguard; A, B, D improve methodological quality, not ethics
17	C	Unfalsifiability — failures are explained away, so no test could ever refute the claim
18	B	Constant offset in one direction = systematic; scatter about the true value = random
19	C	A confounding variable (summer beach attendance) can drive both series; A and B assert unsupported causation
20	D	B: $90m > 60$ m in dollars, but A: 60/1200 = 5% > B: 90/3000 = 3% of budget

Section II — marking notes

Q21. The linear model runs question → hypothesis → experiment → conclusion. Priestley's investigation *began* with an unexpected observation (the mint "restoring" the air) made without a prior hypothesis about plants; the hypothesis (plants restore air) was formulated afterwards and then tested with candle and mouse trials — observation preceding hypothesis, and the investigation redirecting as results emerged. 1 mark for the linear model, 1 for the observation-first departure, 1 for linking to his subsequent testing.

Q22. Detection technologies made invisible radiation observable and measurable (e.g. photographic plates revealed emissions from uranium; counters allowed decay to be quantified), providing the evidence that atoms are not indivisible and have an internal structure that can change — advancing the atomic model. That understanding of radiation and decay enabled a named medical technology, e.g. cobalt-60 radiotherapy or technetium-99m diagnostic imaging. 1 mark each: technology → evidence; evidence → refined atomic understanding; understanding → NAMED medical technology. A response that merely names instruments without the causal chain caps at 1.

Q23. Value to the scientific community: the knowledge identifies which plant, which part and which use are worth investigating, saving years of screening — it directed scientists to a fruit with exceptional vitamin C/antioxidant properties (2 marks). Ethical partnership: free, prior and informed consent from the knowledge holders; recognition of Aboriginal Peoples as the source of the knowledge; and negotiated benefit-sharing (economic returns, joint ownership/IP arrangements) rather than uncompensated commercialisation (2 marks). Conflating the scientific community's perspective with Indigenous perspectives — the 2024 flagged fault — caps at 2.

Q24. Pseudoscience presents claims in scientific-sounding language without the method behind them. The data shows non-reproducibility: four practitioners applying the same "diagnostic technique" to the same 20 photographs selected 3, 11, 7 and 16 photographs respectively, with no two selecting the same set — a genuine diagnostic method should give consistent results between trained practitioners. Iridology also proposes no testable mechanism linking iris patterns to kidney function, and its claims are typically not submitted to controlled trial or peer review. Marks: definition-level feature of pseudoscience (1), explicit use of the supplied figures (1–2), reproducibility/testability reasoning (1). Answers that assert "no evidence" without engaging the data cap at 2.

Q25. The graph shows a correlation, not causation: both variables plausibly depend on a third — suburb population (or suburb size). Larger suburbs have more streetlights AND more residents, hence more recorded asthma cases, so the relationship can exist with no causal link between lights and asthma. Full marks require: correlation ≠ causation stated (1), the confounder named (1), and the mechanism through which it produces both trends (1). (Comparing asthma *rates per person* would remove the confounder — creditable extension.)

Q26. Sample answer: IV — type of watering solution (diluted seaweed extract vs plain water); DV — seedling growth measured quantitatively as height in mm (or change in height) every 2 days for 3 weeks. Take 20 bean seedlings of the same species, age and initial height; randomly assign 10 to receive 50 mL of diluted seaweed extract daily and 10 (the experimental control) to receive 50 mL of plain water daily. Keep constant: light exposure, temperature, soil type and volume, pot size, watering

volume and time. Reliability: the 10 seedlings per group act as repeats — average the growth per group; repeat the whole investigation. Marking (checklist): IV and DV with quantitative measurement (1), water-only control group distinct from controlled variables (1), two named controlled variables (1), sample size with random assignment (1), repetition/averaging for reliability (1), logical sequenced method that would actually work (1). "Repeat it" offered as a validity feature earns nothing for the control criterion.

Q27 (a). Exclude 55.6 g (inconsistent with the other four digital readings); mean of remainder = (49.9 + 50.1 + 50.0 + 50.0)/4 = **50.0 g**.

(b). The digital balance: its mean (50.0 g) equals the true value of the standard mass, while the spring balance's mean (48.04 ≈ 48.0 g) is about 2.0 g below 50.00 g. Accuracy must be judged against the TRUE value, with both means quoted.

(c). Systematic error — readings are consistently offset in one direction (all ≈ 2 g low). Likely cause: a zero/calibration error (e.g. the spring balance not zeroed, or a stretched spring). Correction: re-zero/recalibrate against the standard mass (or subtract the constant offset). "Human error" scores zero.

Q28 (a). Any two justified, e.g.: require prior safety evidence (in-vitro/animal or dermal toxicity data) BECAUSE the cream is untested on humans and first exposure carries unknown risk; require informed consent detailing possible side effects and the right to withdraw at any time BECAUSE participants must understand the risk they accept; require monitoring and a stopping rule if adverse reactions occur. 1 mark per recommendation + 1 for justifications tied to THIS trial's risk (untested topical product). Recommendations about sample size or instrument accuracy address validity, not ethics — no marks (the 2024-flagged confusion).

(b). E.g. the Declaration of Helsinki — requires that research on humans be preceded by risk assessment, that informed consent be obtained, and that the participant's welfare take precedence over the interests of science; this directly addresses the consent/safety issue in (a). Nuremberg Code or NHMRC National Statement equally creditable. 1 mark for the named instrument, 1 for what it requires (not a narrative of historical breaches).

Q29. Two distorting features must be named from the graph: (1) the truncated vertical axis (starting at 94, not 0) makes a rise of about 3 complaints (~3%) look like a doubling in bar height; (2) the "2026 (projected)" bar presents an invented future value (104) alongside measured data, implying an accelerating trend no measurement supports. Effect: readers conclude tap-water quality is rapidly worsening, steering them toward bottled water — the company's commercial interest. Marks: each feature identified with its distorting mechanism (1 + 1), effect on public perception (1), link to the sponsor's interest / explicit judgement (1). Generic "the graph is misleading" without the named features caps at 1.

Q30. Caution because the funder has a commercial interest in a favourable outcome, which can shape research through question selection, study design, selective reporting or suppression of unfavourable results — even at a university, funding can bias what is published (publication bias), so independent replication is needed. Supporting example from a DIFFERENT industry, e.g. the tobacco industry funding research that downplayed the link between smoking and cancer for decades, or fossil-fuel-

funded contrarian climate research. Marks: mechanism of funder influence (2), declared-funding nuance (declaration enables scrutiny but does not remove the bias) (1), specific different-industry example (1). An energy-drink or general food example — same industry — misses the final mark, per the 2025 marking demand.

Q31 (a). Increase = 35.7 − 34.0 = 1.7 kg; percentage = 1.7/34.0 × 100 = **5.0%**, well below the claimed 15% — the trial's own data does not support the claim. Both the calculation and the explicit comparison to 15% required.

(b). All participants simultaneously began a new supervised weight-training program — a confounding variable that could fully explain the gain; with no control group training WITHOUT the bar, the effect of the bar cannot be separated from the effect of the training.

(c). E.g. add a control group drawn from the same gym doing the identical training program but eating no bar (or a placebo bar); any observed difference between groups can then be attributed to the bar — this isolates the IV and makes the design valid. (Random allocation/blinding also creditable if justified via isolating the bar's effect.)

Q32 (a). Not valid because the comparison does not isolate the music: the music group was on a windowsill (light, warmth) while the silent group was in a dark cupboard, so light is a confounding variable changed along with the IV; additionally the pots differed in size and soil volume (a second uncontrolled variable). Two specific faults for 2 marks; "no repeats" is a reliability fault and does not earn validity marks here.

(b). Three modifications, each tied to its criterion, e.g.: place BOTH groups in identical light/temperature locations, with the silent group behind a sound barrier — improves VALIDITY (only the music now differs); use identical pots, soil type and volume — improves VALIDITY (removes further confounds) [accept as a second validity fix]; run more plants per group and repeat the investigation, averaging growth — improves RELIABILITY (consistency of results); measure height with a fixed reference point/digital height gauge, same person measuring at the same time of day, several measurements over time — improves ACCURACY (closeness of each measurement to the true height). 1 mark per modification-plus-named-criterion; a modification linked to the wrong criterion earns nothing.

(c). Fifty plants would improve reliability (more consistent averages) but the light confound would remain: both larger groups would still differ in light as well as music, so the conclusion "music increases growth" would still not follow. Sample size cannot repair a validity fault — the flagged discriminator.

Q33. Recruit runners with knee pain; randomly allocate to two groups. Treatment group wears the magnetic strap; placebo group wears an identical strap with non-magnetic metal inserts of the same mass and appearance. Neither the participants nor the researchers assessing pain know which strap is which (allocation coded by a third party). Pain measured quantitatively (e.g. standard pain scale before/after 6 weeks). Justifications: the placebo controls for the expectation of relief — any psychological benefit occurs equally in both groups, so a real magnetic effect appears as a between-group difference (2 with specification of the placebo); blinding participants prevents expectation bias in self-reported pain, and blinding researchers prevents unconscious bias in measuring/recording (2

with who-is-blinded specified); random allocation distributes other differences between groups (1). Definitions without deployment in THIS scenario cap at 2.

Q34 (a). Marking: axes correct (mass of sugar on x as IV, gas volume on y) with labels and units (1); uniform scales using most of the grid (1); all six points plotted accurately (1); a smooth CURVE of best fit that rises and flattens near 50 mL, drawn ignoring the outlier at (8, 75) (1). A straight line forced through all points, or a line drawn through the outlier, loses the final mark — the planted trap is double: the trend is non-linear AND one point is anomalous (the noted loose bung justifies exclusion).

(b). Gas volume increases steeply with sugar mass at first, then levels off at about 50 mL beyond ~6–8 g. The hypothesis is supported only up to that point: beyond it, added sugar produces little extra gas (the yeast, not the sugar, becomes the limiting factor), so "always increases" is not supported. Trend description quoting values (1); judgement with the plateau reasoning (1).

Q35 (a). Total funding grew from $16.7 b to $26.7 b, but its composition reversed: government's share fell from 55% ($9.2 b) to 38% ($10.1 b) while business's share rose from 40% ($6.6 b) to 57% ($15.3 b) — business is now the dominant funder. Both the share reversal and quoted values required; noting that government funding rose in dollars while falling as a share (the proportion trap) distinguishes top answers.

(b). Direction: business funding flows toward research with commercial application, so applied/product-focused research expands while curiosity-driven pure research (reliant on the shrinking government share) is squeezed; direction increasingly set by market priorities rather than scientific or public-interest priorities. Timeframes: commercial funders favour short development cycles and deliverables; long-horizon research (decades-long studies, fundamental physics) depends on stable public funding and becomes harder to sustain. Reporting: commercial funders may delay publication for patents, report favourable results selectively, or impose confidentiality — pressuring research integrity — whereas public funding generally requires open publication. Judgement: e.g. the growth in total funding benefits science's capacity, but the shift concentrates influence in funders with interests in outcomes; a balanced verdict explicitly weighing benefit against risk earns the final mark. Marks: one per developed strand (direction, timeframe, reporting) = 3, use of the stimulus contrast (1), explicit supported judgement (1).

Q36. Top-band response analyses regulation's effectiveness using all three views: Student A — international instruments exist (e.g. the Declaration of Helsinki; UNESCO's declarations on the human genome; the Nuremberg Code's legacy) and set shared norms, but most lack binding enforcement, supporting A's aim while exposing its practical limit (a named instrument required). Student B — regulatory asymmetry is real ("ethics dumping"/research moving to permissive jurisdictions, e.g. the widely condemned 2018 gene-edited-babies case proceeded outside effective oversight), yet B's alternative is weak alone: the same case shows training and publication norms did not prevent the work — they enabled its condemnation afterwards. Student C — institutional ethics committees (in Australia, HRECs operating under the NHMRC National Statement) are the layer that actually reviews each project prospectively, but they inherit their standards from national/international codes, so C's "only practical safeguard" understates the system's nesting. Judgement: effective regulation is multi-layered — international norms + national law + institutional review — no single layer suffices. Marks: engagement with each statement including its weakness (3), named

code/regulator accurately deployed (2), accurate supporting example/case (1), coherent overall judgement on the extent of effectiveness (2). Criteria mirror the 2023–2025 convention: a response that ignores one student's statement caps below the top band.

Prediction provenance

Which consensus prediction each part of this paper realises. Full evidence panels: intuitioneducation.com.au/hsc-2026

Paper item	Prediction	Agreement
Section I Q1	`topic-level consensus`	— · 2 (opus, grok)
Section I Q2	`topic-level consensus`	0.49 (gpt-5.6-sol) · 2 (gpt-5.6-sol, fable)
Section I Q3	`topic-level consensus`	0.85 (fable) · 2 (fable, opus)
Section I Q4	`watch list`	0.85 (fable), 0.65 (opus) · 2
Section I Q5	`topic-level consensus`	0.70 (grok), 0.65 (fable) · 2
Section I Q6	`topic-level consensus`	0.65 (grok), 0.65 (opus) · 3 (grok, opus, deepseek)
Section I Q7	`topic-level consensus`	0.80 (opus) · 2 (opus, grok)
Section I Q8	`topic-level consensus`	0.76 (gpt-5.6-sol) · 2 (gpt-5.6-sol, opus)
Section I Q9	`topic-level consensus`	0.70 (gemini-3.1-pro) · 2 (gemini-3.1-pro, opus)
Section I Q10	`watch list`	0.75 (grok) · 3 (grok, fable, opus)
Section I Q11	`watch list`	0.35 (deepseek) · 1
Section I Q12	`topic-level consensus`	0.60 (fable), 0.67 (gpt-5.6-sol) · 3
Section I Q13	`watch list`	0.85 (fable), 0.55 (gpt-5.6-sol) · 3 (fable, gpt-5.6-sol, opus)
Section I Q14	`watch list`	0.35–0.60 · 2 (fable, grok)
Section I Q15	`watch list`	0.80 (fable) · 2 (fable, opus)
Section I Q16	`topic-level consensus`	0.57 (gpt-5.6-sol), 0.55 (opus) · 2
Section I Q17	`watch list`	0.58 (gpt-5.6-sol) · 2 (gpt-5.6-sol, fable)
Section I Q18	`topic-level consensus`	0.70 (fable) · 3 (fable, opus, deepseek)
Section I Q19	`topic-level consensus`	0.50 (deepseek) · 2 (deepseek, gpt-5.6-sol)
Section I Q20	`topic-level consensus`	0.52 (gpt-5.6-sol) · 2 (gpt-5.6-sol, opus)
Q21	`is-q9-named-scientist-linear-model`	0.59 · 6 (all)
Q22	`watch list`	0.52 panel mean; 0.35 (opus bold) · 6 (all)
Q23	`is-q10-bioharvesting`	0.59 · 5 (fable, opus, grok, gemini-3.1-pro, deepseek)
Q24	`watch list`	0.35–0.60 · 4 (fable, gpt-5.6-sol, gemini-3.1-pro, deepseek)
Q25	`is-q11-correlation-causation`	0.58 · 5 (fable, opus, gpt-5.6-sol, gemini-3.1-pro, deepseek)
Q26	`is-q2-write-valid-method`	0.70 · 4 (fable, opus, grok, deepseek)
Q27	`is-q5-instrument-error-comparison`	0.62 · 5 (fable, opus, gpt-5.6-sol, grok, deepseek)
Q28	`is-q8-regulation-ethics`	0.60 · 6 (all)
Q29	`is-q12-media-misrepresentation`	0.53 · 5 (fable, opus, gpt-5.6-sol, grok, deepseek)
Q30	`watch list`	0.70 (fable), 0.55 (opus) · 2 (grok rests it)

Paper item	Prediction	Agreement
Q31	`is-q4-claims-numeric-chain`	0.64 · 5 (fable, opus, gpt-5.6-sol, grok, deepseek)
Q32	`is-q3-flawed-design-critique`	0.69 · 5 (fable, opus, gpt-5.6-sol, gemini-3.1-pro, deepseek)
Q33	`is-q6-placebo-double-blind`	0.60 · 5 (fable, opus, gemini-3.1-pro, grok, deepseek)
Q34	`is-q1-graph-construction-trap`	0.78 · 6 (all)
Q35	`is-q7-funding-influence`	0.60 · 6 (all)
Q36	`watch list`	0.42–0.50 · 3 (fable, opus, gpt-5.6-sol)

Built by an AI panel and backtested against the hidden 2025 papers before publication — method at intuitioneducation.com.au/hsc-2026/how-we-did-it. Scored publicly in November 2026. Independent Intuition Education material from publicly available NESA documents; not affiliated with or endorsed by NESA.