

Research assistant: study guide

Science, laboratory, agriculture, food, and environment · suite apt-201-research-assistant · generated 2026-09-15T15:27:03.012Z

Detail	Value
Title	Research assistant: study guide
Generated	2026-09-15T15:27:03.012Z
Fixture/version	apt-201-research-assistant
Sector	Science, laboratory, agriculture, food, and environment
Guide version	v2

Private practice result. Not an official exam certificate, employer decision, hiring signal, admissions decision, or guaranteed outcome. Scores stay on this device unless you export them.

Answer keys and scoring logic stay server-side and are never included in any download or export.

How to use this guide

This file contains the whole study outline for this suite: every skill it draws on, the full lesson for each of those skills, worked examples, practice tips, a glossary, and where each piece of material comes from. Nothing here is a summary of a page you still have to visit.

1. Read the skills section end to end once, without timing yourself.
2. Work the guided practice mode for the suite, using the practice tips as a checklist.
3. Move to the mini-test only when guided practice feels unhurried.
4. Sit the full simulation last, once, in the conditions you expect on the day.

Practice attempts are stored on the device you used, never on an account. Exporting a result is the only way anything leaves that device.

Practice modes and durations

Practice modes for Research assistant

Mode	Duration	What it is for
Guided practice	15 minutes	Untimed, with feedback after every item.
Mini-test	18 minutes	A short timed set for checking pace.
Full simulation	45 minutes	Full length and full time, in one sitting.

- Start guided practice:
<https://learn.novusstreamsolutions.com/aptitude/practice/new?bank=apt-201-research-assistant&mode=guided>
- Start mini-test:
<https://learn.novusstreamsolutions.com/aptitude/practice/new?bank=apt-201-research-assistant&mode=mini>
- Start full simulation:
<https://learn.novusstreamsolutions.com/aptitude/practice/new?bank=apt-201-research-assistant&mode=full>

Skills covered, in full

This suite draws on 5 skill constructs. Each one below carries its complete lesson.

Attention and concentration

Sustained focus, selective attention, and accuracy under time pressure.

Attention and concentration is the skill of still noticing on minute forty of a checking block as reliably as you did on minute two, and of pointing the noticing at the right thing when two things compete. It has three distinguishable parts - selective attention, sustained attention or vigilance, and divided attention - and assessments load them differently: cancellation and checking tasks load vigilance, conflict tasks load selection, and dispatch-style monitoring loads division. It is the construct that decides whether a records clerk, a dispatcher, an air-side controller or a quality inspector catches the one wrong digit in a shift. Practice here is device-local, with no account and nothing uploaded.

What you should be able to do after this lesson:

1. Count occurrences of a target in a dense string without double-counting or losing your place, using a fixed scan path rather than free looking.
2. Separate hits, misses and false alarms in your own results and work out whether you are set too eager or too cautious for the scoring rule in force.
3. Predict where in a long block your error rate will rise, and schedule micro-resets against that curve instead of pushing through it.
4. Recognise a transposition error, the class that survives a same-characters gist check and is therefore missed most often.
5. Explain why an automatic process such as reading interferes with a controlled one, and use that to anticipate which distractors will actually cost you time.
6. Run a two-stream monitoring task by alternating on a fixed cadence rather than attempting genuine simultaneity.

Worked examples and pitfalls

A cancellation count you can actually check: Count every 7 in this row: 4 7 1 7 7 3 9 7 2 8 7 5 7 6 0 7.

Working left to right and tapping once per hit: hits at the second, fourth, fifth, eighth, eleventh, thirteenth and sixteenth positions. That is seven sevens. Two error modes produce nearly all the wrong answers. The first is the adjacent pair - the 7 7 at positions four and five gets counted once, because the eye takes a repeated character as one perceptual object. The second is losing the place after the 9, where the visually similar 9 and 7 force a moment of re-checking and the scan restarts a character early, producing an over-count of eight. The defence for both is a fixed scan path with a physical anchor: a fingertip or cursor moving one character at a time, never jumping back. Re-scanning to check is the thing that creates the double-count, so if you must verify, verify by counting a second time from the RIGHT-hand end and comparing totals, never by re-reading part of the row.

Hits, misses and false alarms: the eager candidate loses: A checking batch contains 300 records, 40 of which are genuinely faulty. Candidate A flags 46 records and 36 of them are truly faulty. Hits 36, misses 40 minus 36 equals 4, false alarms 46 minus 36 equals 10. Candidate B flags 38 and 34 are truly faulty: hits 34, misses 6, false alarms 4. On hit rate alone A wins, 36 out of 40 which is 90 percent against B's 34 out of 40 which is 85 percent. Now apply the scoring rule that most checking tasks actually use, where a false alarm cancels a hit: A scores 36 minus 10 equals 26, B scores 34 minus 4 equals 30. B wins by four despite catching two fewer faults. This is why 'flag anything that looks odd' is bad advice on a scored checking task, and why the first thing to read on a checking item is whether wrong flags are penalised. Your response criterion - how much evidence you demand before flagging - is adjustable, and it should be set from the

scoring rule, not from your temperament.

The transposition that passes every gist check: Compare these two lines and decide whether they match. Invoice 4820-7391-06. Invoice 4820-7931-06. They do not: the middle group reads 7391 in the first and 7931 in the second, with the 3 and the 9 swapped. Transpositions are the most-missed error class in record checking for a structural reason - the character SET is identical, the length is identical, the first and last characters of the group are identical, so every fast check the visual system runs comes back clean. Substitutions and omissions change the character inventory and get caught; transpositions do not. Two habits raise the catch rate. Read digits in fixed groups of two rather than as a whole number, so 73-91 against 79-31 becomes a mismatch at the first group instead of a subtle difference somewhere in a four-digit blur. And check groups in a deliberately non-natural order - last group, first group, middle group - because reading left to right lets the confirmation you built at the start carry you through the middle, which is exactly where the swap is usually planted.

Conflict: why reading fights you: The classic demonstration is Stroop's, published in 1935: the word RED printed in blue ink, with the instruction to name the ink colour. Naming takes measurably longer, and errors go up, because reading a familiar word is automatic and cannot be switched off, so the automatic response has to be suppressed before the controlled one can be produced. The same conflict has a numeric version you can test on yourself in a second: how many characters are in the string 4 4 4? The answer is three, and the digit 4 pulls at you the entire way. In an assessment this appears wherever the salient feature and the asked-for feature come apart - a chart where the tallest bar is not the answer to the question, a form where the highlighted field is not the one being verified, a row where the bold total is not what the stem requested. The practical move is to name the target feature out loud before you look - 'ink colour', 'character count', 'the value for March' - because pre-loading the target biases selection before the automatic reading response gets a chance to win.

Where the errors actually appear in a 45-minute block: Errors in a long checking block are not spread evenly. Mackworth's 1948 clock-watching study established the pattern that gives the effect its name: detection declines over a prolonged watch, with the sharpest deterioration early rather than at the very end. Practically, on a 45-minute self-timed checking block, expect your per-minute error rate in minutes 20 to 45 to run visibly above minutes 1 to 20 even though nothing about the material changed, and expect the subjective sense of effort to lag the actual decline, so it will not feel like you are getting worse. Two things work against it. Break the block into three fifteen-minute segments with a ten-second reset between them - look away, unfocus, breathe out - which costs thirty seconds of a 45-minute block, about one percent of the time, and buys back more than that in caught errors. And score your practice by segment rather than as one number, because a single overall accuracy figure hides exactly the information you need: whether your problem is skill, which shows as flat error rate, or endurance, which shows as a rising one.

Two streams: alternate on a cadence, do not try to merge: A dispatch-style monitoring task: keep a running total of the numbers announced on channel one while watching channel two for the code word AMBER. Genuine simultaneity is not available - the two tasks compete for the same control resource - so the choice is not whether to alternate but whether to alternate deliberately or accidentally. Deliberate looks like this: fix the arithmetic to a rhythm, updating the total only at each announcement and holding a single number between updates, which frees the gaps for channel two. If the total is 34 and the next announcement is 7, you spend under a second reaching 41 and then you are free again. Accidental looks like re-deriving the running total from the beginning because you did not trust it, which locks up the whole window and is when AMBER goes past unheard. The measurable failure of divided attention is almost never a failure to hear the target; it is a failure to have any spare capacity at the moment it arrived. The related phenomenon worth knowing is inattention blindness, illustrated by Simons and Chabris in 1999: an unexpected and perfectly visible event is missed entirely when attention is committed to a counting task.

How to practise this skill

- Fix your scan path before you start, and never re-scan backwards to double-check. Backward re-scanning is what produces both double-counts and lost places; if you need to verify a count, run it again from the other end and compare.
- Read the scoring rule before the first item. Whether false alarms are penalised changes the correct strategy completely, and a criterion set for the wrong rule costs more marks than slow reading does.
- Check numeric strings in fixed groups of two or three and in a deliberately shuffled group order. Transpositions are the errors that survive whole-string comparison, and they are the ones planted on purpose.
- Time your practice in segments and log accuracy per segment, not per session. A flat error rate means you need speed; a rising one means you need endurance and breaks, and the two problems have opposite fixes.
- Name the target feature out loud before each conflict item. Pre-loading the relevant dimension is the cheapest available defence against the salient-but-wrong answer.
- Practise with an actual distractor present rather than in silence. A test hall has movement, coughing and page-turning, and attention practice in perfect quiet trains a condition you will not meet.

Glossary

Selective attention: Prioritising one source or feature while suppressing competing ones. It is what fails when the visually salient element of a display captures the response instead of the requested one.

Sustained attention (vigilance): Maintaining detection performance over a long, low-event period. It is the capacity that checking, monitoring and inspection tasks are really sampling.

Vigilance decrement: The decline in detection performance over a prolonged watch, typically steepest early in the block. Mackworth's 1948 clock test is the standard public reference.

Divided attention: Handling two streams that compete for control. Performance is best modelled as fast alternation with a switch cost, not as genuine parallelism.

Hit, miss, false alarm, correct rejection: The four outcomes of any detection decision: flagging a real fault, missing one, flagging a clean record, and correctly leaving a clean record alone. Any honest accuracy claim needs at least the first three.

Response criterion: How much evidence you require before flagging. Shifting it trades misses against false alarms without changing your underlying sensitivity, which is why it must be set from the scoring rule.

Stroop interference: The slowing that occurs when an automatic response, such as reading a word, conflicts with the requested response, such as naming its ink colour. Named for Stroop's 1935 experiments.

Inattentional blindness: Failing to notice a fully visible, unexpected event because attention is committed elsewhere - the effect Simons and Chabris demonstrated in 1999 with a counting task.

Study this next

- Attention and concentration skill hub:
<https://learn.novusstreamsolutions.com/aptitude/skills/attention-concentration>
- Cognitive Skills Lab: attention practice:
<https://learn.novusstreamsolutions.com/cognitive-skills/attention-concentration>
- Emergency dispatch and 911 communications suite:
<https://learn.novusstreamsolutions.com/aptitude/tests/emergency-dispatch-and-911-communications>
- Error detection lesson: <https://learn.novusstreamsolutions.com/aptitude/skills/error-detection/lesson>
- Perceptual speed lesson: <https://learn.novusstreamsolutions.com/aptitude/skills/perceptual-speed/lesson>
- Attention and concentration lesson on Novus Learn:
<https://learn.novusstreamsolutions.com/aptitude/skills/attention-concentration/lesson>

Where this material comes from

- Character strings, invoice comparisons, batch counts and monitoring scenarios above are written for Novus Learn. All figures are original and no published test item is reproduced.
- Terminology and directions of effect follow widely published work: Mackworth (1948) on the vigilance decrement, Stroop (1935) on response conflict, Simons and Chabris (1999) on inattention blindness, and the standard hit/miss/false-alarm framing from signal detection theory. Concepts only; no result figures are attributed to those studies.
- Novus Learn aptitude construct registry (catalog seed) for construct scope and suite mapping.
- Public educational framing only - not affiliated with any official exam board, publisher or employer, and not a clinical, diagnostic or attention-disorder screening measure.

Educational preparation only. Novus Learn does not administer official exams and does not guarantee scores or hiring outcomes.

Data interpretation

Tables, charts, graphs, dashboards, trends, comparisons, and evidence-based conclusions.

Data interpretation is the discipline of getting a correct number out of a table, chart or dashboard that was not built to make your question easy, and of saying so when the data cannot answer it at all. It dominates graduate and analyst screening, and it is the section where strong arithmetic still fails, because the marks are lost in the header row, the axis scale and the wording of the question. The same skill is the daily work of anyone who reports on a management pack, a clinical audit or a stock ledger. Practice is stored on this device only; there is no account and nothing is uploaded unless you export it.

What you should be able to do after this lesson:

1. Read a value correctly out of a table or chart including its units multiplier, footnotes and any 'excludes' or 'provisional' qualifier attached to the row.
2. Compute share of total, percentage change and percentage-point change from the same pair of cells, identify which of the three a question is asking for, and separate a movement in a rate from a movement in the underlying count.
3. Work with index numbers relative to a base year, including why a change of five index points is almost never a five percent change.
4. Join two tables on a shared key and produce a normalised figure (per head, per unit, per thousand) rather than comparing raw totals.
5. Recognise chart presentation effects (truncated axes, dual axes, cumulative versus periodic series), and answer from the numbers rather than from the visual impression.
6. Apply the 'cannot say' discipline: state precisely which extra fact would be needed before the question becomes answerable.

Worked examples and pitfalls

Read the header: (£000) changes every answer by a factor of a thousand: Table titled 'Regional revenue, year to March (£000)': North 1,240; South 986; East 1,455; West 719. Total = $1,240 + 986 + 1,455 + 719 = 4,400$, so the business turned over 4,400 thousand pounds, that is 4.4 million. East's share is $1,455 / 4,400 = 33.1$ percent. The whole set: North 28.2 percent, South 22.4, East 33.1, West 16.3, summing to 100. Two things go wrong here. The first is reading East's revenue as 1,455 pounds and then reporting a business with a total turnover of 4,400 pounds, which nobody notices because every option is scaled the same way, until the question asks for revenue in millions and only one option is right. The second is the comparison wording. East's share is $(1,455 - 719) / 4,400 = 16.7$ percentage points above West's, and East's revenue is $(1,455 - 719) / 719 = 102$ percent more than West's, that is slightly more than double. 'Sixteen point seven' and 'a

hundred and two' both describe the same two cells honestly, and the question decides which one is correct. Note that the percentage-point figure must be computed from the unrounded shares rather than by subtracting the rounded ones, or the last digit will not survive.

One pair of rows, three correct increases: A complaints table: 2023, 120,000 orders, complaint rate 4.0 percent; 2024, 150,000 orders, complaint rate 5.0 percent. Three defensible answers to 'how much did complaints increase?'. The rate rose by 1.0 percentage point. The rate rose by $(5.0 - 4.0) / 4.0 = 25$ percent in relative terms. And the count of complaints rose from $0.04 \times 120,000 = 4,800$ to $0.05 \times 150,000 = 7,500$, which is $(7,500 - 4,800) / 4,800 = 56.25$ percent. All three are arithmetically right; only one answers the question in front of you. The pattern to internalise is that a rate and a count move together only when the denominator is fixed, and here it is not. Order volume grew 25 percent as well. If the question is about customer experience, the rate is the honest figure; if it is about how many complaint handlers to hire, the count is. Test items usually ask for the one you would not have chosen.

Index numbers: five points is not five percent: A cost index with 2020 = 100 reads 104 in 2021, 111 in 2022 and 109 in 2023. From 2021 to 2023 the index rose 5 points, but the percentage change is $5 / 104 = 4.8$ percent, because the base for the comparison is 104, not 100. From 2022 to 2023 it fell 2 points, which is $-2 / 111 = -1.8$ percent, and note that costs fell even though the index remains 9 percent above the 2020 base. A level and a change are different claims. The only comparison where points and percent coincide is against the base year itself: 2020 to 2023 is 100 to 109, exactly plus 9 percent. Watch also for a rebased series, where a table switches to 2022 = 100 partway down; the two segments cannot be compared directly without converting one of them, and an item that quietly rebases is testing whether you read the column heading.

Joining two tables: totals and per-head figures disagree on purpose: Table 1, headcount by site: Leeds 84, Derby 47, Bristol 129. Table 2, absence days recorded in the same period: Leeds 630, Derby 300, Bristol 903. 'Which site has the worst absence problem?' On raw totals Bristol is worst at 903 days. Normalise per head and the ranking changes: Leeds $630 / 84 = 7.50$ days per employee, Bristol $903 / 129 = 7.00$, Derby $300 / 47 = 6.38$. Leeds is worst, Bristol is merely biggest. The organisation-wide figure is $1,833 / 260 = 7.05$ days per head, which is a useful reference line: Leeds is above it, the other two below. The general rule is that any comparison between units of different size demands a denominator, and the denominator has to come from the other table. Items are built so the raw-total answer and the per-head answer are both on the option list, and so the site with the biggest total is never the site with the highest rate.

The truncated axis: measure the numbers, not the bars: A quarterly satisfaction chart with a y-axis running from 78 to 82 shows bars at 79.2, 79.8, 80.4 and 81.1. Visually the last bar looks several times taller than the first, because only the top 4 points of a 100-point scale are drawn. The actual movement is $81.1 - 79.2 = 1.9$ points, which on the score's own scale is a relative rise of $1.9 / 79.2 = 2.4$ percent. If the question asks 'by approximately what percentage did satisfaction improve', the answer is about 2 percent, and the distractor built from the bar heights will be something like 40 or 400 percent. Related presentation effects to check before answering: a dual-axis chart where two series use different scales and appear to cross meaningfully when they do not; a cumulative series, where a flattening line still means the total is growing, just more slowly; and a logarithmic axis, where equal vertical distances are equal ratios rather than equal amounts. In every case the defence is the same. Find the printed numbers, or read the gridline values, and compute.

Cannot say: revenue is not profit: A product table shows units sold and total revenue. Product P: 4,200 units, 71,400 pounds. Product Q: 1,800 units, 41,400 pounds. Average selling price is $71,400 / 4,200 = 17.00$ for P and $41,400 / 1,800 = 23.00$ for Q, so Q earns more per unit while P earns more in total. Now the statement to evaluate: 'P is more profitable than Q.' The correct response is cannot say. Profit needs cost, and the table has no cost column; a product with a 17 pound price and a 16 pound unit cost is less profitable than one priced at 23 with a cost of 9, and nothing here rules that out. Contrast with 'Q generated more revenue per unit than P', which the table fully supports and which is true. The habit worth building is to finish every cannot-say judgement with the missing input named out loud ('cannot say, because unit cost is not given') because that forces you to distinguish a genuinely unanswerable item from one you simply have not worked

hard enough on.

How to practise this skill

- Read the title, the units line, the row and column headers and any footnote before you look at a single value. Roughly the first fifteen seconds of an item should contain no arithmetic at all, and that fifteen seconds is what prevents the thousand-fold and percentage-point errors.
- For every item, write down which of the three quantities is wanted (share of total, relative change, or change in percentage points), before computing. Most wrong answers on this construct are correct arithmetic applied to the wrong quantity.
- Whenever two groups differ in size, ask what the denominator should be. If a question compares sites, teams, countries or periods of unequal length and you have not divided by something, you are almost certainly answering the wrong question.
- Practise the cannot-say items separately and force yourself to name the missing variable each time. Candidates who train only on computational items reliably over-answer inference statements under time pressure.
- Do not redraw or re-scale charts in your head. Locate the gridline values or the data labels, and if neither exists, interpolate between two labelled gridlines and state the bound rather than guessing a point value.
- Time yourself per item rather than per section. Data interpretation sets share a stimulus, so the first item costs the reading time and the rest should be fast; if item four takes as long as item one, you did not build a mental map of the table.

Glossary

Units multiplier: A scaling note in a table title or column header, such as (£000), (millions) or (per 1,000 population). It applies to every value in scope and is the single most common source of order-of-magnitude errors.

Index number: A series rescaled so a chosen base period equals 100. Changes between two non-base periods must be divided by the earlier value, so a movement in index points is not a percentage change except when measured from the base.

Rebasing: Restating an index against a new base period. Segments of a series with different bases cannot be compared directly, and a table that rebases partway down is testing whether you read the headings.

Truncated axis: A chart whose value axis does not start at zero, which exaggerates the apparent size of differences between bars or points. Legitimate for showing small movements in a large quantity, misleading if read as area or height.

Cumulative series: A line showing a running total rather than each period's value. It can only go up or stay flat, so a flattening cumulative line means the periodic figure is falling, not that the total is.

Weighted average: An average in which each value is multiplied by the size of the group it represents. Averaging two group percentages directly is only correct when the groups are the same size, which in these tables they rarely are.

Normalisation: Dividing a raw figure by an exposure measure (headcount, units sold, population, days open), so groups of different size can be compared. The denominator usually lives in a second table.

Cannot say: The verdict when a statement is neither supported nor contradicted by the data supplied. A correct cannot-say answer can always be defended by naming the specific missing variable.

Study this next

- Data interpretation skill hub: <https://learn.novusstreamsolutions.com/aptitude/skills/data-interpretation>
- Data interpretation practice in the Cognitive Skills Lab:

<https://learn.novusstreamsolutions.com/cognitive-skills/data-interpretation>

- Numerical reasoning lesson:
<https://learn.novusstreamsolutions.com/aptitude/skills/numerical-reasoning/lesson>
- Financial data interpretation suite:
<https://learn.novusstreamsolutions.com/aptitude/tests/financial-data-interpretation>
- Scientific data interpretation suite:
<https://learn.novusstreamsolutions.com/aptitude/tests/scientific-data-interpretation>
- Data interpretation lesson on Novus Learn:
<https://learn.novusstreamsolutions.com/aptitude/skills/data-interpretation/lesson>

Where this material comes from

- Every table, index series and chart described above was constructed for Novus Learn, and each figure was verified by recomputing the totals and the reverse calculation.
- Definitions of index numbers, rebasing and weighted averages cross-checked against standard public references such as the Wikipedia articles 'Index (economics)' and 'Weighted arithmetic mean'. Terminology only; no data or item text is taken from any source.
- Novus Learn aptitude construct registry (catalog seed) for construct scope and suite mapping.
- Public educational framing only: not affiliated with any official exam board, publisher or employer, and no copyrighted test item is reproduced.

Educational preparation only. Novus Learn does not administer official exams and does not guarantee scores or hiring outcomes.

Situational judgment

Evaluating workplace responses against role-relevant principles.

Learn a repeatable way to compare workplace responses: establish the facts, identify duties and risks, respect role boundaries, then choose a proportionate first action. This is educational preparation, not an official scoring guide.

What you should be able to do after this lesson:

1. Separate facts stated in a scenario from assumptions that the scenario does not support.
2. Rank response options by immediate risk, policy or role obligations, proportionality, and follow-through.
3. Explain why a strong first action is better than passive, punitive, or unauthorized alternatives.

Worked examples and pitfalls

Worked scenario: an unverified safety concern: A colleague reports a possible equipment fault while a deadline is approaching. First distinguish the known fact, the report, from the unverified cause. A strong response protects people and affected work, checks the concern through the right channel, tells the relevant lead, and records what was done. Ignoring the report underreacts; shutting down unrelated work or accusing someone before checking the facts overreacts.

Method: facts, duties, risks, response: Write four short notes before ranking options: what is known, who may be affected, which duty or boundary applies, and what safe next step is available now. Prefer an action that addresses the immediate issue and creates useful follow-through. Do not reward an option merely because it sounds decisive.

How to practise this skill

- Answer the question asked: best first action, worst action, or complete response are different tasks.
- Check whether an option acts within the person's authority and escalates only as far as the risk requires.
- When two options look reasonable, prefer the one that gathers missing facts and communicates ownership.

Glossary

Proportionality: Matching the urgency and scope of a response to the evidence, likely impact, and authority available.

Role boundary: The limit of what a person may decide or do without approval, specialist help, or escalation.

Follow-through: Confirming ownership, recording the decision, and checking that the issue was actually resolved.

Study this next

- Situational judgement test:
<https://learn.novusstreamsolutions.com/search?q=Situational%20judgement%20test>
- Workplace ethics: <https://learn.novusstreamsolutions.com/search?q=Workplace%20ethics>
- Decision-making: <https://learn.novusstreamsolutions.com/search?q=Decision-making>
- Situational judgment lesson on Novus Learn:
<https://learn.novusstreamsolutions.com/aptitude/skills/situational-judgement/lesson>

Where this material comes from

- Novus Learn original situational-judgment suite scenarios and published construct mapping.
- Novus educational framework: facts, duties, risks, role boundaries, proportional action, and follow-through.

Educational preparation only. Novus Learn does not administer official exams and does not guarantee scores or hiring outcomes.

Verbal reasoning

Drawing conclusions from written information without relying on outside assumptions.

Verbal reasoning is the discipline of answering strictly from a passage: deciding whether a statement is True, False, or Cannot Say on the evidence given, and refusing to import anything you happen to know about the topic. It is the workhorse construct in graduate and civil-service screening, banking and consulting sifts, and language-focused public-service batteries, where a passage about parcel volumes or grant conditions is followed by four statements to classify. The same discipline is what stops a contract review, a grant assessment or an incident summary from quietly acquiring facts nobody wrote down. Practice here is device-local: no account, and nothing leaves this device unless you export it.

What you should be able to do after this lesson:

1. Classify a statement as True, False, or Cannot Say, and state in one sentence which words in the passage force the classification.
2. Explain the difference that costs the most marks: 'Cannot Say' means undetermined by the passage, while 'False' means contradicted by it.
3. Spot quantifier drift between passage and statement (all, some, most, only, none) and show why 'all A are B' never licenses 'all B are A'.
4. Refuse a causal upgrade: recognise when a passage reports co-occurrence or sequence and a statement

claims cause.

5. Separate what a passage asserts from what it reports somebody else asserting, and classify statements about each correctly.
6. Do the small arithmetic hidden inside verbal items - percentage changes, fractions of a stated total - without leaving the passage for outside data.

Worked examples and pitfalls

The three-way rule on one short passage: Passage: 'The Riverton depot handled 4,200 parcels in March, a 12 percent increase on February. A night shift was introduced in January; in March it processed 40 percent of the parcels the depot handled.' Now take four statements. (1) 'The depot handled 3,750 parcels in February.' February is March divided by 1.12: $4,200 / 1.12 = 3,750$ exactly, so this is True - the passage determines it even though the number is not printed. (2) 'The night shift caused the March increase.' Cannot Say. The passage puts the night shift and the increase in the same depot in the same period; it never claims one produced the other, and it never denies it either. Most candidates mark this False, reasoning correctly that correlation is not causation but then choosing the wrong label: False is reserved for statements the passage contradicts. (3) 'The night shift processed more than 1,700 parcels in March.' Forty percent of 4,200 is 1,680, which is not more than 1,700, so this is False - contradicted by arithmetic on the passage's own figures. (4) 'The depot handled more parcels in March than in January.' Cannot Say: January is mentioned only as the month the shift started, and no January volume is given.

Quantifier drift and illicit conversion: Passage: 'All accredited suppliers submit quarterly audits. Some suppliers in the northern region are accredited.' Statement A: 'Some suppliers in the northern region submit quarterly audits.' True, and it is worth seeing why the chain holds: at least one northern supplier is accredited, and every accredited supplier submits, so at least one northern supplier submits. Statement B: 'Every supplier that submits quarterly audits is accredited.' This reverses the first sentence. 'All accredited suppliers submit' leaves the door wide open for unaccredited suppliers to submit as well - perhaps voluntarily, perhaps under another rule - so the answer is Cannot Say. Turning 'all A are B' into 'all B are A' is called illicit conversion and it is the single most productive trap in this format, because the reversed sentence sounds like a paraphrase. Statement C: 'No unaccredited supplier submits quarterly audits' is the same reversal wearing a negative coat, and it is Cannot Say for the same reason. Statement D: 'All northern suppliers submit quarterly audits' is also Cannot Say: 'some are accredited' says nothing about the rest.

Asserted versus reported: Passage: 'The committee's report claims that the scheme cut waiting times by a third. Two of the four regional boards have disputed that figure.' Statement A: 'The scheme cut waiting times by a third.' Cannot Say. What the passage asserts is that a report claims this; the claim's truth is never settled, and the dispute does not settle it either. Statement B: 'At least two regional boards disagree with the report's waiting-time figure.' True, and note 'at least' - the passage says two disputed it, and two of four disputing is consistent with 'at least two'. Statement C: 'All four regional boards accept the figure.' False, directly contradicted. Statement D: 'The report is wrong.' Cannot Say. This item type appears constantly in policy and diplomacy passages, where a paragraph stacks a claim, a source, and a reaction, and the reader has to keep three different truth-conditions apart. The reliable habit is to underline the reporting verb - claims, argues, estimates, alleges, projects - and remember that everything downstream of it belongs to the source, not to the passage.

Percentages are not symmetric: Passage: 'Membership fell from 8,000 to 6,000 over two years, then recovered by 25 percent in the third year.' Statement: 'Membership at the end of year three was higher than at the start.' The fall is 2,000 on a base of 8,000, which is 25 percent. The recovery is 25 percent of the new base: $6,000 \times 1.25 = 7,500$. That is below 8,000, so the statement is False. The trap is elegant, because a 25 percent fall followed by a 25 percent rise feels like it should cancel. It does not: to get back from 6,000 to 8,000 you need a 33.3 percent rise, since $2,000 / 6,000 = 0.333$. Whenever a verbal item states a change as a percentage, write down what the percentage is a percentage of before you touch it. A related version

supplies the recovery as an absolute number instead - 'recovered by 2,000 members' - which does return the total to 8,000, and candidates who answered the first version from memory get the second one wrong.

Verbal analogies: name the relation as a sentence: Item: 'ANTISEPTIC is to INFECTION as ____ is to ____.' Options: (a) vaccine : immunity, (b) insulation : heat loss, (c) medicine : illness, (d) bandage : wound. Write the stem relation as a full sentence first: 'An antiseptic is applied in order to prevent an infection from occurring.' Now test each option against that exact sentence. Option (a) runs the other way - a vaccine is given to produce immunity, not to prevent it. Option (c) is the right family but the wrong specificity: medicine typically treats an illness that has already started. Option (d) is applied after the wound exists. Option (b) fits precisely: insulation is applied in order to prevent heat loss from occurring. The general method is the same on the simpler item types. 'BOOK is to LIBRARY as PAINTING is to ____' has canvas (the material), artist (the maker), and gallery (the place a collection is kept and shown) among its options; all three are genuine relations to 'painting', and only the third matches the stem sentence. Options in analogy items are chosen to be true statements about the words, just not the stated relation.

Two premises with 'some' prove nothing: Argument: 'Some depot managers are qualified engineers. Some qualified engineers hold a rigging certificate. Therefore some depot managers hold a rigging certificate.' This is invalid, and the way to prove it is a counter-model rather than an argument. Let the depot managers be Ana and Ben; let the qualified engineers be Ben and Cara; let the certificate holders be Cara alone. Premise one holds: Ben is a manager and an engineer. Premise two holds: Cara is an engineer with the certificate. The conclusion fails: neither Ana nor Ben holds the certificate. Since a single consistent world makes both premises true and the conclusion false, the argument cannot be valid, so in a True / False / Cannot Say framing the conclusion is Cannot Say. Two premises that both begin 'some' never combine into a conclusion, because 'some' gives you an overlap without telling you where the overlap sits. Contrast a valid pair: 'All accredited labs are inspected annually. No inspected facility is exempt from the fee.' Every accredited lab is inspected, and no inspected facility is exempt, so no accredited lab is exempt - True.

How to practise this skill

- Answer from the passage even when you know the subject. Candidates with a background in the topic score worse on verbal items than they expect, because their own knowledge quietly supplies the missing premise that turns a Cannot Say into a True.
- Run the two-worlds test on every candidate 'Cannot Say': can you imagine a world where all the passage's sentences hold and the statement is true, and another where they hold and it is false? If both, it is Cannot Say. If only the false world exists, it is False.
- Underline the quantifiers and hedges in the statement (all, some, only, most, may, must, likely) and find their counterparts in the passage. Roughly half of the wrong answers in this construct come from a single word swapped between the two.
- Budget about 25 to 30 seconds per statement rather than per passage, and read the passage once for structure before touching the statements. Re-reading the whole passage for each statement is what causes candidates to run out of time on the last set.
- Keep a wrong-answer log with one label per error: quantifier, causation, reported-versus-asserted, arithmetic, outside knowledge. A clear pattern almost always emerges within about forty items, and it is usually a single label.
- Work untimed until the three-way rule is automatic, then add the clock. Attempts are stored on this device only, so a slow first pass through a passage set costs you nothing but the time you spend on it.

Glossary

Entailment: A statement is entailed by a passage when it cannot be false while every sentence of the passage is true. Entailment is the only thing that earns a 'True' in this format; being plausible, likely, or well known does not.

Cannot Say: The verdict for a statement the passage neither entails nor contradicts. It is a claim about the passage, not about the world, which is why a statement you know to be true in real life can still be Cannot Say.

Quantifier: A word fixing how much of a group a claim covers: all, most, some, few, none, only. Swapping one quantifier for another changes the logical content completely while barely changing how the sentence reads.

Illicit conversion: The invalid move from 'all A are B' to 'all B are A', or from 'if P then Q' to 'if Q then P'. It preserves the words and destroys the logic, which is why converted sentences make such effective wrong answers.

Counter-model: A concrete, consistent scenario in which the premises hold and the conclusion fails. Producing one is the fastest possible proof that an argument is invalid, and it takes two or three named individuals.

Hedge: A qualifier such as may, could, is expected to, or is associated with. A hedged sentence in a passage cannot support an unhedged statement, and an unhedged passage sentence is not weakened by a hedged statement.

Reporting verb: A verb such as claims, argues, estimates, or alleges that attributes the following content to a source. Everything downstream of it is the source's assertion, and the passage takes no position on it.

Analogy relation: The specific link between the two stem words in an analogy item - tool and user, item and container, action and purpose. Naming it as a full sentence before reading the options removes almost all of the guesswork.

Study this next

- Verbal reasoning skill hub: <https://learn.novusstreamsolutions.com/aptitude/skills/verbal-reasoning>
- Practise the verbal reasoning bank:
<https://learn.novusstreamsolutions.com/cognitive-skills/verbal-reasoning>
- Critical thinking lesson: <https://learn.novusstreamsolutions.com/aptitude/skills/critical-thinking/lesson>
- Reading comprehension lesson:
<https://learn.novusstreamsolutions.com/aptitude/skills/reading-comprehension/lesson>
- General civil-service aptitude suite:
<https://learn.novusstreamsolutions.com/aptitude/tests/general-civil-service-aptitude>
- Verbal reasoning lesson on Novus Learn:
<https://learn.novusstreamsolutions.com/aptitude/skills/verbal-reasoning/lesson>

Where this material comes from

- Worked items written for Novus Learn. Every passage, statement set and figure above is original and invented for this lesson; no published or copyrighted test item is reproduced.
- Terminology follows standard, widely published usage in introductory logic and assessment writing - entailment, quantifier, illicit conversion, counter-model.
- Novus Learn aptitude construct registry (catalog seed) for the construct scope and the suite mapping shown in the related links.
- Public educational framing only - not affiliated with any official exam board, publisher or employer.

Educational preparation only. Novus Learn does not administer official exams and does not guarantee scores or hiring outcomes.

Critical thinking

Evaluating evidence, assumptions, arguments, credibility, and alternative explanations.

Critical thinking assessments ask you to take an argument apart: what is being concluded, what it rests on, what evidence would settle it, and whether a proposed conclusion actually follows. Published critical-thinking batteries typically run five task types - inference, recognition of assumptions, deduction, interpretation, and evaluation of arguments - and they are deliberately built so that agreeing with a conclusion and judging the argument as strong come apart. The construct is heavily weighted in policy analysis, audit, legal, investigative and graduate selection, and it is the same skill that stops a plausible chart from turning into a bad decision. Practice stays on this device; there is no account and nothing is uploaded unless you export it.

What you should be able to do after this lesson:

1. State an argument's conclusion in your own words before evaluating it, and identify which sentences are premises and which are background.
2. Use the negation test to separate a required assumption from a statement that would merely strengthen the argument.
3. Distinguish validity from truth, and identify affirming the consequent as distinct from the valid modus tollens form.
4. Compute a posterior probability on a screening example and explain why the rate of true positives among all positives is far lower than intuition suggests.
5. Generate at least two alternative explanations - selection, reverse causation, a common third factor - for any claimed causal effect, and name the comparison that would rule them out.
6. Judge argument strength on relevance and directness rather than on agreement, including marking arguments you personally reject as strong.

Worked examples and pitfalls

The negation test finds assumptions; nothing else does: Argument: 'The council should fit bin sensors across the district. In the trial depot they cut collection trips by a fifth.' Candidate assumption A: 'The trial depot's waste pattern is broadly representative of the rest of the district.' Negate it - suppose the trial depot is nothing like the rest of the district - and the argument collapses, because a result that does not transfer supports nothing about the district. So A is a required assumption. Candidate assumption B: 'Bin sensors are the cheapest available technology for this purpose.' Negate it - suppose they are the most expensive - and the argument still stands as given, since it argued from a reduction in trips, not from cost. B is not required; it would strengthen a cost-based argument that was never made. Candidate C: 'Fewer collection trips reduce total operating cost.' Negate it and the recommendation loses its point, so C is required too, even though the argument never mentions cost. That last case is the one candidates miss, because the required assumption is what bridges the evidence to the recommendation, and bridging assumptions are by definition unstated. The failure mode across this whole item type is picking statements that support the conclusion instead of statements the argument cannot survive without.

Base rates: 90 percent accurate, 15 percent right: A production line runs an automated defect test. Defects occur in 1 percent of units. The test flags 90 percent of genuinely defective units, and wrongly flags 5 percent of good units. A unit has just been flagged - how likely is it to be defective? Work in whole units rather than probabilities. Take 10,000 units: 100 are defective and 9,900 are good. Of the 100 defective, the test flags 90. Of the 9,900 good, it wrongly flags 5 percent, which is 495. Total flags: $90 + 495 = 585$. Of those, 90 are genuine, so the answer is $90 / 585 = 15.4$ percent. Better than the 1 percent base rate, and nowhere near the 90 percent that intuition offers. The error has a name - confusing the probability of a flag given a defect with the probability of a defect given a flag - and it has a practical consequence: the rework queue has to be sized for 585 units a batch, not 100, and 495 of the 585 units sitting in it are perfectly good. Whenever a test, screening rule or model accuracy figure appears in an item, ask how many of the negatives there are,

because with a rare condition the false positives from a large clean population swamp the true positives from a small affected one.

Three alternative explanations, and the comparison that settles it: Claim: 'Employees who attended the optional resilience workshop took 30 percent fewer sick days last year, so the workshop works.' Alternative one, selection: attendance was optional, so the people who signed up may have been the healthier and more engaged to begin with, and would have taken fewer sick days regardless. Alternative two, reverse direction: employees who were frequently unwell were the least able to attend a full-day workshop, so illness determined attendance rather than the other way round. Alternative three, a common third factor: the workshop ran on Thursday afternoons at head office, so attendance largely tracks being an office-based rather than shift-based worker, and shift workers take more sick leave for reasons that have nothing to do with resilience training. What would settle it: randomly assigning the offer to half the workforce and comparing the two groups. Failing that, the cheapest useful check already exists in the payroll data - compare attendees' and non-attendees' sick days in the year before the workshop. If attendees were already 30 percent lower, selection is doing the entire job. The habit worth building is to name the comparison group before you accept any before-and-after number.

Deduction: affirming the consequent versus modus tollens: Rule: 'If an invoice is over 10,000 euro, it requires two signatures.' Argument A: 'This invoice has two signatures, therefore it is over 10,000 euro.' Invalid - this is affirming the consequent. The rule makes two signatures necessary for large invoices; it says nothing that prevents a cautious manager from double-signing a 400 euro invoice, so a two-signature invoice of any size is consistent with the rule. Argument B: 'This invoice has only one signature, therefore it is not over 10,000 euro.' Valid - this is modus tollens, denying the consequent to deny the antecedent, and it holds assuming the rule was followed. Argument C: 'This invoice is not over 10,000 euro, therefore it does not have two signatures.' Invalid again, denying the antecedent. Two further points that deduction items test directly. First, validity is about form alone: 'All banks close on Sundays; Riverton Mutual is a bank; therefore Riverton Mutual closes on Sundays' is perfectly valid even if the first premise is false, and a valid argument with a false premise can deliver a false conclusion. Second, in these items you must accept the premises as given even when you know them to be untrue, because the question is whether the conclusion follows, not whether it is true.

Interpretation: what a survey figure does and does not license: Data: all 1,200 employees of a firm were surveyed. 62 percent said they would accept a four-day week at 90 percent pay. Of those who said yes, 71 percent were under 35. Conclusion one: 'A majority of this firm's employees would accept the trade.' This follows - 62 percent of a complete census of the firm is a majority, and the conclusion is properly limited to this firm. Conclusion two: 'A majority of the firm's under-35 employees would accept the trade.' This does not follow, and the arithmetic shows why. Sixty-two percent of 1,200 is 744 accepters, and 71 percent of 744 is about 528 accepters aged under 35. That is the composition of the accepters, not the acceptance rate within the under-35 group, and without knowing how many under-35s the firm employs the rate is undetermined: if there are 900, the rate is $528 / 900 = 59$ percent; if there are 600, it is 88 percent; if there were 1,100 the rate would be below half. Conclusion three: 'The firm should move to a four-day week.' This does not follow either - a stated preference about pay says nothing about output, shift cover, customer hours or cost. Notice that conclusion two is the base-rate error from the screening example, wearing survey clothing.

Strong and weak arguments on the same question: Question: should the agency publish raw inspection scores for every site? Strong argument for: 'In the neighbouring authority, publication was followed by a measurable fall in repeat violations, which is the stated aim of the inspection regime.' It is directly about this decision, it addresses the regime's own objective, and it offers evidence - and if true it would change your view. Weak argument for: 'Yes, because transparency is always good.' Sweeping, unsupported, and it would apply identically to publishing anything at all, which is the tell. Strong argument against: 'Raw scores omit the severity weighting, so a site with one critical failure can rank above a site with six minor ones, and a member of the public cannot see the difference.' Specific, mechanistic, directly about the proposal, and it would

change your view. Weak argument against: 'No, because businesses will object.' It may be true and it may matter politically, but it does not bear on whether publication achieves the regime's aim, and no evidence is offered. Two tests separate strong from weak: is the argument about the exact question asked, and would accepting it change the decision? Note that you can hold a firm personal view on publication and still be required to mark the argument on the other side as the strong one - that split is what the item type is measuring.

How to practise this skill

- Negate every candidate assumption out loud. If the argument survives the negation, it was never an assumption, however supportive it sounds. This single habit converts the assumption item type from guesswork into a mechanical check.
- Write the conclusion in your own words before reading any option. Half of the wrong answers on inference and evaluation items are responses to a conclusion the argument never reached.
- Ask 'percentage of what?' on every percentage in the stimulus and reconstruct the denominator in whole units. Base-rate and composition errors both dissolve the moment you write out counts instead of rates.
- Deliberately practise marking arguments you disagree with as strong and arguments you agree with as weak. Assessments in this construct are built to catch agreement masquerading as evaluation, and the effect is largest on politically loaded stimuli.
- For any causal claim, write two alternatives - selection and a third factor - and name the comparison group that would rule them out, before you decide whether the evidence supports the claim.
- Log wrong answers by task type rather than by topic: inference, assumption, deduction, interpretation, evaluation. Candidates are rarely weak across all five, and the profile tells you where the next hour of practice belongs.

Glossary

Conclusion: The claim an argument is trying to establish. It is not always last, is often signalled by therefore, so, or should, and locating it correctly determines every subsequent judgement about the argument.

Assumption: An unstated premise the argument requires in order to work. Identified by negation: negate it and a genuine assumption brings the argument down, while a merely helpful statement leaves it standing.

Validity and soundness: An argument is valid when the conclusion cannot be false while the premises are true, which is a property of form alone. It is sound when it is valid and the premises are actually true.

Affirming the consequent: The invalid pattern 'if P then Q; Q; therefore P'. It is the most common deductive distractor because it differs from the valid modus tollens form by only the position of a negation.

Base rate: How common something is in the population before any test or evidence is applied. Ignoring it makes accurate-sounding tests appear far more informative than they are when the condition is rare.

Confounder: A third factor associated with both the supposed cause and the outcome, capable of producing the entire observed relationship on its own. Ruled out by randomisation or by an explicit comparison group.

Selection effect: A difference between groups created by how people entered them rather than by the treatment under study. Optional programmes and voluntary surveys are where it appears most often.

Falsifiability: The property of a claim that some observation could show it to be wrong. A claim compatible with every possible result carries no information, which is why 'what would change your mind?' is a diagnostic question.

Study this next

- Critical thinking skill hub: <https://learn.novusstreamsolutions.com/aptitude/skills/critical-thinking>
- Practise the critical thinking bank: <https://learn.novusstreamsolutions.com/cognitive-skills/critical-thinking>

- Verbal reasoning lesson: <https://learn.novusstreamsolutions.com/aptitude/skills/verbal-reasoning/lesson>
- Government audit suite: <https://learn.novusstreamsolutions.com/aptitude/tests/government-audit>
- Investigator and detective reasoning suite:
<https://learn.novusstreamsolutions.com/aptitude/tests/investigator-and-detective-reasoning>
- Critical thinking lesson on Novus Learn:
<https://learn.novusstreamsolutions.com/aptitude/skills/critical-thinking/lesson>

Where this material comes from

- Worked items written for Novus Learn. The arguments, survey figures, defect-test numbers and evaluation options above are original and invented for this lesson; no published or copyrighted test item is reproduced.
- The five task types named in the summary - inference, assumption, deduction, interpretation, evaluation of arguments - describe a structure used publicly across several critical-thinking assessments; no affiliation with any publisher is claimed or implied.
- Terminology follows standard, widely published usage in introductory logic and research methods - validity, soundness, affirming the consequent, base rate, confounder, selection effect.
- Novus Learn aptitude construct registry (catalog seed) for the construct scope and the suite mapping shown in the related links.
- Public educational framing only - not affiliated with any official exam board, publisher or employer.

Educational preparation only. Novus Learn does not administer official exams and does not guarantee scores or hiring outcomes.

Recommended preparation

- Practice attention and accuracy:
<https://learn.novusstreamsolutions.com/aptitude/skills/attention-concentration/lesson>
- Build data interpretation: <https://learn.novusstreamsolutions.com/aptitude/skills/data-interpretation/lesson>
- Practice situational judgement:
<https://learn.novusstreamsolutions.com/aptitude/skills/situational-judgement/lesson>

Answer keys, scoring and privacy

Answer keys and scoring logic stay server-side and are never included in any download or export.

Formal suite answer keys and scoring logic stay on the server and are not part of any download, in any format. Downloadable keys exist only for the open, untimed practice material (the practice packs on the puzzles, cognitive-skills and reasoning practice lab pages), where the answers are already public teaching content.

Novus Learn needs no account. Your practice history lives in this browser's storage on this device, and is sent to a server only if you create an account and switch on backup. This file contains no attempt, session or result link, so it is safe to share.

Private practice result. Not an official exam certificate, employer decision, hiring signal, admissions decision, or guaranteed outcome. Scores stay on this device unless you export them.

Answer keys and scoring logic stay server-side and are never included in any download or export.

Source: <https://learn.novusstreamsolutions.com/aptitude/tests/research-assistant/study-outline>