

Reward detective

Ages 9–13 with an adult • About 20–30 minutes • Free trial edition

Your role: Help with reading, let the learner predict before running, and ask for evidence before showing the explanation below. No coding knowledge is required. Shorten the activity or take a break when needed.

Before starting

Open wovi.world/agent-maze. Check **Low -2** and **Off +0**, especially after using a shared build link. Ask: "What does the agent score? What do you think it will do?" The learner must select a prediction to unlock the first run.

What to look for in the first run

On the current map with those settings, the agent hits the storm on its second move. At that choice, going east earns **+4 progress +2 energy -2 storm = +4**. Going north also scores **+4**. It is a tie, and the fixed scan order checks east first. The storm route does not have a higher score.

Change one setting, compare the result

Switch only Safety to **High -12**; leave Memory **Off +0**. At the same choice, east now scores **+4 +2 -12 = -6** while north remains **+4**. In the current map, this setting reaches the beacon in 10 moves without a storm collision. Ask the learner to compare their actual run with this explanation.

What counts as a useful explanation?

Look for a link between the changed rule and observed behavior: "The energy no longer cancels the storm penalty, so the safe move scores higher." Accept age-appropriate wording. The learner should also notice that a person changed the setting. This is a fixed scoring simulation; the agent is not learning a new policy from repeated runs.

Warehouse question: one possible answer

Guard: Exclude no-entry routes before scoring speed; if every route is blocked, stop for a person. **Test:** Make the quickest urgent-parcel route cross the no-entry area while a slower permitted route exists. Check that the robot takes the permitted route. A further test can block every route and check that it stops.

Other proposals are discussion opportunities. Ask whether the rule merely discourages entering the area or actually rules it out, and what observation would show it failed. A larger penalty alone is not a guarantee in every possible situation.

Keep a small family record

Keep the learner's prediction, changed setting, observed result and explanation on paper. No name or upload is needed. This record helps your conversation; it is not a score, certificate or proof of mastery.

This is Session 1 of the complete free trial edition. Keep all four worksheets together. The existing games remain free.

Example explorer

Ages 9–13 with an adult · About 20–30 minutes · Free trial edition

Purpose: Help the learner connect a prediction to its closest training example, distinguish a false alarm from a missed risk, and notice the limits of reusing the same test cases. Read the expected results after the learner predicts.

Prepare the exact comparison

Start a fresh Reef Rescue mission at wovi.world/demo. Reload the game tab if it is mid-mission; this restarts the current simulation. The three fixed examples are Healthy coral garden, Plastic fragments and Polluted runoff. For the first scan, select **Heat-stressed coral** as the one optional example, making four. Other first choices can produce different results.

Expected results for these selections on the current four cases		
Check	First scan: 4 examples	Second scan: all 6
Correct classifications	3 / 4 (75%)	4 / 4 (100%)
False alarms	1	0
Missed risks	0	0

Read the evidence, not just the score

In the first scan, **Cloudy sand plume** is predicted RISK although the ranger label is SAFE: a false alarm. Its closest example is **Polluted runoff**. Hidden warm pocket and Floating bottle are correctly labeled RISK; Seaweed nursery is correctly labeled SAFE. Seaweed nursery initially matches Healthy coral garden, so the missing seaweed example does not cause an error in this particular test.

What changes in the second scan?

After **Review mistakes & improve Wovi**, add both **Natural seaweed meadow** and **Passing sand current**, leaving the original four selected. The game enables scanning at six examples. Cloudy sand plume now matches Passing sand current and is labeled SAFE. Seaweed nursery now matches Natural seaweed meadow and remains SAFE. The other two labels stay RISK; no classification errors remain on these four cases.

Explain that this model assigns each test case the label of its nearest example. The visible closest match helps explain the changed sand-plume decision. Because this game run adds two examples at once, do not present the overall improvement as an experiment that changed only one example.

Support the learner’s explanation and next test

A useful answer is: “The first set treated harmless sand like polluted runoff. With a reviewed sand-current example, the sand plume gets a closer SAFE match.” In the game’s final question, **Wovi saw more diverse evidence** is the matching explanation. The world and four test cases stayed the same.

For the paper test plan, keep the first four examples, classifier and four test cases fixed; add only Passing sand current and inspect the sand-plume label and its nearest example. This describes a comparison to run in a tool that permits five examples; it is not an extra scan available in the current game. Ask for another situation too, such as heat stress combined with cloudy water, without promising how the model would classify it.

State the limit clearly

These are four fixed, hand-designed game cases. After seeing mistakes and changing the examples, scanning those same cases again is practice and debugging, not an independent test on unseen data. The displayed percentage is not real-ocean accuracy. “Match” is a distance-based display, not a calibrated probability that the answer is correct. New cases would be needed to investigate whether the rule generalizes.

Keep: the first and second results, the changed closest example, a one-example test plan and one limitation. The family keeps these paper notes; they are not uploaded or automatically assessed.

Transfer lab

Ages 9–13 with an adult • About 20–30 minutes • Free trial edition

Your role: let the learner make a prediction before checking these notes. Read the short cases together and help with arithmetic if needed; the useful work is explaining a choice with evidence.

1. Greenhouse: a penalty and a hard rule do different things

Reference scores for the three successive rules			
Plan	Start	Update 1	Update 2
A	$1 + 2 \times 5 = 11$	$11 - 3 = 8$	$8 + 4 = 12$
B	$2 + 2 \times 4 = 10$	10	10
C	$3 + 2 \times 3 = 9$	9	9

The choices are **A → B → A**. Update 1 discourages giving seedlings too little water. Update 2 keeps that penalty, but its new bonus outweighs it. The bonus applies only to A because only A gives flowers 5 cups.

With the hard rule “seedlings get at least 2 cups,” remove A first. B then beats C, **10 to 9**, under Update 2. Ask: “Did we change the points, or change which plans were allowed?”

A larger penalty could also select B for these exact numbers; that is a reasonable numerical proposal. Ask the learner to test it against a still larger bonus. A hard limit better expresses a requirement that must hold regardless of rewards. If a future plan list has no permitted choice, the helper needs a “no valid plan; ask the gardener” response instead of silently dropping the rule.

Further tests: a larger flower bonus; an A-like plan with even fewer seedling cups; or a list with no plan meeting the minimum. Reward choices can vary when the gardener’s preferences vary, but do not silently change the stated minimum.

2. Leaf photos: a total can hide a weaker group

The total is **20 correct out of 25 (80%)**. Bright photos: **18/20 (90%)**. Dim photos: **2/5 (40%)**. A drawing that shows those fractions is enough; percentage calculation is optional. The larger bright group makes the overall result look stronger than the dim-group result.

B is the most direct next experiment: add varied dim training photos of both leaf types. A may improve bright-light coverage but does not directly fill the missing lighting condition. C reuses known test answers; a better score on those same photos would not show that unfamiliar photos are handled better.

Keep fresh bright and dim photos separate from training, then compare both groups again. A learner may suggest balancing the new test groups to make the comparison easier; that is reasonable if the group sizes are reported. The five dim tests show a problem worth investigating, not that every dim photo will fail or that a particular improvement is guaranteed.

The table does not establish why the helper missed those photos. Dim examples are a sensible hypothesis to test because they were missing from training; label mistakes or other photo differences are also worth checking. Accept an alternative if the learner names the evidence they would need.

3. Museum labels: a big score is not a reason to skip every check

A: gap $8 - 3 = 5$ and the image is clear, so use “Cup” as a temporary label under the stated rule. **B:** gap $6 - 5 = 1$, so request review. **C:** gap $9 - 2 = 7$, but the image is blurred, so request review anyway. “OR” means either reason is enough.

For C, asking for a clearer image is a useful first action. For B, inspecting the object may add evidence that the photo lacks. Leaving either unlabelled while evidence is unavailable is also acceptable. A person should examine evidence, not merely repeat the top guess.

A useful boundary test is a **clear image with points 7 and 4**: the gap is exactly 3, so the rule allows a temporary top label. Blurring the same image should switch the result to review. A clear tie should also request review.

The rule sorts cases for this paper exercise; the points do not prove any label is correct. Ask the learner what they would do if the temporary label were later found wrong. Revising it and reviewing the rule is more useful than treating the original score as proof.

Close with one comparison

Look for a distinction such as: "Extra training photos do not change the greenhouse's allowed plans," or "A larger label score does not repair a blurred image." Let the learner use their own words. Keep the prediction, decision, evidence and proposed next test; these are discussion notes, not a grade.

Wovi AI Investigator · Session 3 · Parent reference · wovi.world/family-pack-session-3

Explain & improve

Ages 9–13 with an adult · About 20–30 minutes · Free trial edition

Your role: help the learner turn their own work into an explanation. A small, accurate record is more useful than a polished report with missing evidence. Allow a drawing or spoken answer, and write down their words if that helps.

Start from the work that exists

Gather Sessions 1–3. Mark unfinished activities honestly; do not require all of them before discussing one completed investigation. If no activity has been tried, return to an earlier session first. If a prediction was never recorded, leave it as “not recorded”; a new run can supply a new prediction, not reconstruct the old one.

Help the learner choose a specific moment: the changed storm penalty in Agent Maze, one training-example choice in Reef Rescue, or one Transfer lab calculation. “I finished the game” alone does not explain what changed or why.

Check the link between change, observation and explanation

Ask the learner to point to an actual score, count, route or paper calculation. If several settings changed at once, help them say that; do not force a one-change explanation that the record cannot support. They can propose a clearer comparison as their next test.

For reference, Session 3’s greenhouse choices are $A \rightarrow B \rightarrow A$ as the scoring rules change, then B when the minimum-water rule removes A. The leaf test has 20/25 correct overall, with 18/20 in bright light and 2/5 in dim light. Museum cases B and C need review for different reasons. Use these only if the learner chose those cases; they are not evidence that your child completed them.

If using the Agent Maze sample, distinguish a person changing a fixed scoring rule from an agent learning a new rule by itself. If using Reef Rescue, discuss the examples and displayed result the learner actually observed; do not supply a fictional before-and-after improvement.

Listen first; choose one follow-up

Let the learner give their short tour before correcting it. Choose one follow-up from the learner page and leave time to think. “I would need a new photo to check that” or “I changed two settings, so I cannot tell which mattered” can be useful answers.

If the explanation does not match the evidence, ask them to look again at the relevant number or rule. Help them revise one sentence in their own words. A correct prediction is not required; being able to revise an explanation is a worthwhile outcome of this activity.

Make the next test concrete

A usable plan names what changes, what stays the same and what result would challenge the explanation. For example, a learner who chose the museum task could keep a 3-point gap and compare a clear image with a blurred one. They should predict the two decisions before applying the review rule.

For a missing-example question, a fresh set kept out of training is a better next check than repeatedly testing the examples just taught. For a reward question, a changed reward can test whether a proposed penalty still works. Accept other tests if the learner can explain what they would learn from either outcome.

Use the self-check to choose a next step

“Shown” means the learner can point to that part of today’s record or explain it to you. “Not yet” identifies a useful next step; it is not a failed grade. Do not turn these marks into a mastery score or compare them with another child’s record.

Finish by asking which investigation they would revisit and why. Keep the notes at home. The worksheet does not require a name, upload, public sharing or an automatically generated parent report.