DRAFT FOR COMMENT · FREE FOR ANY SCHOOL TO COPY

Is it worth the *licence?*

A one-term check for SENCOs deciding whether an AI or edtech tool did what it claimed, for the pupils in front of them.



1 afternoon
to set it up

10 minutes
a fortnight to run it

1 hour
at the end to decide

Aevum FOUNDRY·AI

hello@aevumfoundry.ai

01

Who this is for

This is for a SENCO who has been offered an AI or edtech tool, has one free afternoon to set up a trial, and needs to know by the end of term whether to pay for it.

It takes about three hours to set up, ten minutes a fortnight to run, and an hour at the end to decide. No research training needed.

It answers one question: did this tool do what it claimed, for the pupils in front of you, at a cost you can justify? It will not give you an effect size or anything you could publish. It will give you a defensible decision and a record of how you reached it, which is more than most schools have when the renewal invoice arrives.

> **Declared interest.** Aevum Foundry AI makes AI tools for schools, and we would rather be measured with a check like this one than bought on a demo. Nothing in this guide favours our products or anyone else's. If a vendor objects to being tested this way, that tells you something.

02

Why the vendor's evidence won't answer your question

A vendor's study reports an average across hundreds of pupils. Your decision is about three or four children with a particular profile, and the pupils a tool helps most are the ones whose needs are least alike. A tool that lifts the average can do nothing for a child with a working-memory difficulty and a great deal for a child with dyslexia, or the reverse. The average hides exactly the thing you need to know.

The second problem is adoption. When a tool works in one school and not another, most of the difference is usually how it was used, not what it is. The teacher who sets it up carefully, uses it three times a week and checks the output gets a different result from the one who was handed a login in September. If your trial doesn't record how the tool was actually used, you can't tell a product that failed from a product that was never really tried.

So this check does two things the vendor's study can't. It measures named pupils against their own starting point, and it records use alongside outcome so you can separate the two.

03

The afternoon: six steps to set it up

Do these in order. Each has a box on the record sheet at the end.

1. **Write down the claim.** One sentence, in the vendor's words if possible: "This tool will do X for pupils with Y." If you can't get a sentence that specific from the vendor, that is your first finding. Vague claims can't fail, which is why vendors like them.

2. **Pick the pupils.** Three to six, each with a named need the claim is supposed to address. Not your whole caseload. Write the need next to each name, in the terms your school already uses (EHCP outcome, SEN support plan target, reading age, whatever you track). If you can, pick pupils with at least two different profiles, so you learn who it works for and not only whether it works.

3. **Pick one thing per pupil that should change.** It must be something you already measure or can measure in ten minutes: a reading-fluency score, words written independently in a set task, number of prompts needed to start work, homework returned, time to complete a task. One thing per pupil. Resist the urge to track five.

4. **Record the starting point.** Take that measure for each pupil before anyone touches the tool, and date it. This is the step schools skip, and without it nothing at the end means anything. Where you can, take it twice a week apart, so you know how much it bounces on its own.

5. **Decide who uses it, how often, and write it down.** Which adult, which lessons, how many sessions a week, for how long. This is your "dose". When the trial ends and the result is flat, the first question is whether the dose was delivered, and you can only answer that if you set one.

6. **Set the stop rule and note the version.** Agree now what would make you stop early (see the red flags) so the decision isn't made in the heat of a bad week. Then write down the tool's version number or the date you started, and ask the vendor to tell you about any update during the term. A tool that changes halfway through is two tools.

> One more thing before you leave the room. Tell the pupils and their parents what you are trying and why, in plain terms. It takes a letter, and it protects you later.

04

During the term: the ten-minute log

Every two weeks, fill in one row per pupil. Ten minutes, no more. If it takes longer, you are measuring too much.

WHAT TO RECORD	HOW
Sessions actually delivered	Count against the dose you set. Three planned, one delivered, write one.
The measure	Take the same measure as the starting point, same way, same conditions.
Pupil's reaction	One word from the pupil if they can give it: liked, tolerated, refused, upset.
Adult effort	Minutes of adult time the tool cost or saved that fortnight, roughly.
Anything that changed	Staff absence, a new TA, a version update, the pupil moved class.

Two rules. Never skip a fortnight because nothing happened; "nothing happened" is data. And never adjust the measure halfway through because the numbers look wrong. If the measure was a bad choice, note that, keep taking it, and add a better one alongside.

At half term, read the log once. If sessions delivered are under half the dose, fix the delivery before you judge the product.

05

Red flags that end the trial early

Stop the trial, write down why, and tell the vendor if any of these happen. You don't owe a product a full term.

- A pupil is distressed by it more than once, or refuses it. Their refusal is a result.
- The tool gives a pupil wrong information with confidence (a wrong spelling, a wrong method, a made-up fact) and you can't switch that behaviour off.
- You find it collecting data you didn't agree to, or you can't get a plain answer about where pupil data goes.
- A version update changes what the tool does and the vendor didn't tell you.
- It needs more adult time to supervise than the task took without it.
- The vendor asks you for a quote, a case study or a logo before the trial has ended.

06

End of term: making the decision

Sit down with the log and answer four questions, in this order. Write the answers on the record sheet.

1. **Was the dose delivered?** Add up sessions delivered against sessions planned. Under half, and you have learned about your school's capacity to run this tool, not about the tool. That is still a useful answer: a tool you can't deliver is not worth the licence either, whatever it does in the vendor's study.

2. **For each pupil, did the measure move, and by more than it bounced at the start?** Compare the last reading with the starting point. If you took the baseline twice, you know roughly how much it wobbles on its own. A change smaller than that wobble is not a change. Write one of three words per pupil: better, same, worse.

3. **Who did it work for?** Lay the pupils' needs next to their results. If every "better" has one profile and every "same" another, that is the finding, and it is more useful than any average. It tells you who to buy it for, and who not to.

4. **What did it cost in adult time?** Total the effort column. A tool that helps two pupils and costs a TA three hours a week has a price the licence fee doesn't show.

Then decide, using this table.

DOSE DELIVERED	PUPILS BETTER	DECISION
Yes	Most, across profiles	Buy. Record the profiles it worked for.
Yes	Some, one clear profile	Buy for that profile only. Negotiate a smaller licence.
Yes	Few or none	Don't buy. You have a documented reason.
No	Any	Don't judge the tool yet. Decide whether you can fix the delivery and run one more term, or drop it because you can't.

Whatever you decide, keep the record sheet. The next person in your job will be offered the same tool.

07

Seven questions to put to the vendor before you start

Send these in writing and keep the answers with the record sheet. A good vendor will answer all seven in a page. A vague answer to any of them is a finding.

1. In one sentence, what will this do, for which pupils, and how would we know?
2. What evidence do you have, who collected it, and were they paid by you?
3. Which pupil data does it store, where, for how long, and who can see it?
4. Will you tell us before any update during our trial, and can we stay on the current version until the term ends?
5. Can we see what the tool is showing or saying to each pupil, after the fact?
6. What does a school that got nothing out of it look like, and why did that happen?
7. If we buy for three pupils rather than the whole cohort, what does that cost?

> Treat usage statistics (logins, minutes, "engagement") as information about adoption, never as evidence of learning. A pupil can spend an hour a week in a tool and learn nothing from it.

08

What this check is not

It is not a controlled trial. With four pupils and no comparison group you can't rule out that they would have improved anyway, and you shouldn't claim otherwise to governors or to the vendor. What you can say is that you set a claim, measured it, recorded how the tool was used, and made a decision you can defend. That is the standard most purchasing never reaches.

If a tool passes this check and you are thinking of rolling it out across a trust, that is the point to ask for a proper evaluation. The Education Endowment Foundation and the Department for Education are running an EdTech Testbed Programme from 2026 to 2030, with rapid evaluations of six to twelve weeks and longer studies, and the EEF's formal guidance on evaluating AI and edtech tools is due in spring 2027. Use this check to decide which tools are worth that effort.

This guide may be copied and adapted by any school without permission. If you improve it, we would like to see the improved version: hello@aevumfoundry.ai.

The cover photograph is AI-generated. It shows no real school, pupil or product.

RECORD SHEET

Trial header, pupils and measures

Pupil names go on the school copy only; use initials or codes on anything that leaves the building.

TRIAL HEADER

Tool and version, or start date

The claim, in one sentence

Dose: who, which lessons, sessions per week, minutes each

Stop rule

Vendor's written answers received (date)

Parents informed (date)

PUPILS AND MEASURES

PUPIL	NAMED NEED	MEASURE	BASELINE 1 (DATE)	BASELINE 2 (DATE)	WOBBLE
A					
B					
C					
D					
E					
F					

RECORD SHEET

Fortnightly log and end of term

FORTNIGHTLY LOG (ONE ROW PER PUPIL PER FORTNIGHT)

DATE	PUPIL	SESSIONS DELIVERED / PLANNED	MEASURE	PUPIL REACTION	ADULT MINUTES	ANYTHING THAT CHANGED

END OF TERM

Dose delivered (total / planned)

Pupil A: better / same / worse

Pupil B: better / same / worse

Pupil C: better / same / worse

Pupil D: better / same / worse

Which profiles did it work for?

Total adult minutes over the term

Decision and reason

Signed, date