

The 10-Minute AI Tool Test

Updated July 26, 2026

This is the test I run before any AI tool earns a spot in my workflow. One real task, ten minutes on the timer, five scores. A tool that cannot produce useful work in ten minutes on a trial will not suddenly get useful in month three on a paid plan.

The 10-minute test

Do steps 1 and 2 before you start the clock. Everything else happens inside the ten minutes.

- ☐ **1. Pick a real task.** Pull something from your current project, not a toy prompt. A module outline, a scenario draft, a voiceover script, a quiz bank.
- ☐ **2. Sanitize your inputs.** Use real project content that is safe to share. Swap in placeholder names for anything confidential before it goes into a trial account.
- ☐ **3. Start the trial and the timer.** Ten minutes, no pausing.
- ☐ **4. Skip the tour.** Close the onboarding, the templates gallery, and the sample project. Go straight to producing output.
- ☐ **5. Push for work quality.** The goal is output you could show a stakeholder, not a demo you could show a friend.
- ☐ **6. Stop at ten minutes.** Whatever you have on screen is what you score. No "one more prompt."
- ☐ **7. Estimate edit time.** Look at the output and estimate the minutes needed to make it shippable. Compare that to how long building it by hand would take.
- ☐ **8. Check the exit.** Confirm the output actually moves into your authoring tool or LMS: embed, export, or clean copy into your working file.
- ☐ **9. Score it below.** Fill in all five rows, total it, and apply the thresholds.

The scorecard

Score each row from 1 to 5. Be strict. Trials are designed to feel better than daily use.

Criterion	What a 5 looks like	Score (1-5)
Output quality	A usable draft that needs minor edits. A 1 is output you would throw away and rebuild from scratch.	
Edit time vs. hand-build time	Fixing the output is clearly faster than building the thing by hand. A 1 means hand-building wins.	
Workflow fit	Output embeds or exports into your authoring tool or LMS with no rework. A 1 means retyping everything.	

Data and privacy	Clear, published policy on training and retention that would pass your org's review. A 1 is no policy you can find.
Price per useful output	Monthly price divided by outputs you would actually ship is an easy yes. A 1 means you pay for shelf-ware.
Total (out of 25)	

Pass or fail

20 or higher, keep it and put it into real project work this week. 14 to 19, run the test once more on a different real task before deciding. 13 or lower, drop it and cancel the trial today. A data and privacy score of 1 or 2 is an automatic drop no matter what the total says.

Category worksheet

Log every tool you test so you stop retrying tools you already dropped. Two slots per category; print this page again if you test more.

Category	Tool tested	Total (0-25)	Verdict (Keep/Drop)
Analysis			
Analysis			
Authoring			
Authoring			
Media (image, audio, video)			
Media (image, audio, video)			
Simulation and roleplay			
Simulation and roleplay			
Dev and automation			
Dev and automation			
Assessment			
Assessment			

Tip

test one category at a time. Comparing two authoring tools on the same task tells you far more than testing an authoring tool against a media tool.

Have any questions? Email me at devlin@devlinpeck.com