CODERPAD CUSTOMER GUIDE • AI IN THE PAD

Size for the interview you should be running

Your AI package decides what a candidate is able to do in front of you. At the entry level they can question an agent, review a plan, ask it to check the edge cases on a stock-price problem. The agent never writes a line. Further up the ladder they hand over the keyboard and it **builds against a real repository**, running tests and reading the workspace back on every step. Most engineering roles are hiring for that second thing.

Four levels of AI fluency, and what each one lets you see

What kind of interview you can run comes down to the package rather than the plan. Packages move on their own, so a team can change package without touching plan, interview volume or features.

L0 — AI-ASSISTED

STANDARD

Ask-only screens

The candidate questions the agent, reviews a plan, gets an explanation. The agent never writes or runs anything. You come away knowing how somebody reasons, and very little about how they work in a codebase.

L1 — AGENTIC

FRONTIER LITE · FRONTIER

A single agent on a real repo

The agent works against a real repository instead of answering in the abstract. This is the floor for a coding interview that tells you anything. A candidate has to brief it properly and catch mistakes before they spread.

L2 — AI COLLA-BORATION

RECOMMENDED FOR MOST ENGINEERING HIRING

FRONTIER PRO

A full hour of real build work

Multi-file changes, tests, agents running alone or side by side. A candidate can explore the codebase, take a wrong turn, back out of it and still finish inside the hour.

L3 — ORCHES-TRATION

FRONTIER MAX

Orchestration as the skill being tested

Either the build is bigger than an hour, or the thing being assessed is how somebody splits work across several agents. Keep it for multi-day take-homes and sessions built around orchestration.

Level 2

Frontier Pro is where most engineering hiring should sit

A dense hour of real build work, one agent or several

START HERE •

Work out what you want to watch a candidate do. If the answer is that you want to see them **build with AI**, the exercise has to let the agent write code.

TEST IT AGAINST YOUR BEST •

Reaching the ceiling does not end an interview. The session carries on with a lighter model and the interviewer is told. The switch lands wherever the candidate happens to be.

THE DIRECTION OF TRAVEL •

Interviews keep getting more agentic. Whatever a team needs this year is close to the least they will need next year.

What each package unlocks

Each column here is a kind of conversation you can have with a candidate, so read it that way before you read the specs.

Frontier Pro is the package most engineering hiring should be on.

PACKAGE	STANDARD	FRONTIER LITE	FRONTIER	• RECOMMENDED FRONTIER PRO	FRONTIER MAX
The interview it unlocks What a candidate can actually do	LEVEL 0 Ask-only screens	LEVEL 1 A single agent on a real repo	LEVEL 1 A single agent, more room to run	LEVEL 2 A full hour of real build work	LEVEL 3 Multi-hour and orchestration
What you can test The signal it produces	How they reason about code	Whether they can direct an agent	Whether they can direct an agent	How they build, verify and decide	How they decompose work across agents
Included with Packages move independently of plan	Starter, Team	Team Plus	Enterprise	Upgrade	Upgrade
Models available Frontier = Opus-class, flagship GPT, Gemini Pro	Up to Sonnet-class	All models	All models	All models	All models
Claude Code & Codex Real CLI agents running inside the pad	—	✓	✓	✓	✓
Agents at once Running side by side in one session	2	2	3	4	4
Orchestrate agents Directing agents as the assessed skill	—	—	—	✓	✓
Session length it suits Where the format stops fitting	Short screens	Up to an hour	Up to an hour	A dense full hour	Multi-hour, multi-day

What happens when a session reaches the ceiling

The interview carries on. The AI moves to a lighter model and **the interviewer is told the moment it happens**, so nobody is guessing later about what the candidate had to work with. Where the switch lands is the problem. On a small package it tends to arrive while somebody is still building, in the stretch of the hour you most wanted to watch.

What you can actually test at each level

A candidate can ask an agent questions for an hour without the session growing much. Signal starts when they hand over the keyboard and the agent begins writing and running code. **Level 2 is the first package where that holds up for a full hour.**

LEVEL 0

Standard

AI-assisted: the candidate asks, the agent answers

The candidate walks through an approach and asks the agent to check the edge cases on it. Take a problem like finding the largest loss in a stock price series. You can cover a lot of ground this way and learn real things about how somebody reasons. You will not see them working inside a codebase.

- Shows how somebody judges an approach the AI has put in front of them
- Leaves the question of whether they can build with AI completely open
- Suits a narrow, early-stage screen

LEVEL 1

Frontier Lite
Frontier

Agentic: one agent, writing and running code

The agent works against a real repository. A candidate has to brief it properly, read what comes back, and catch the mistakes before they spread through the rest of the work. Single-agent build sessions can run long enough to reach this ceiling inside the hour.

- Shows whether somebody briefs an agent properly or throws a sentence at it
- You can see how much of the output they actually read before accepting it
- Frontier runs three agents at once where Frontier Lite runs two

LEVEL 2

Frontier Pro

RECOMMENDED

AI collaboration: a dense hour of real build work

Multi-file changes, tests, agents running alone or side by side. A candidate can explore the codebase, take a wrong turn, back out of it and still finish. The number of agents matters less here than how much building you ask for.

- Multi-file builds where tests are expected as part of the work
- Codebase-comprehension exercises, where somebody explains the system before touching it
- Room for one agent or several across a full interview hour

LEVEL 3

Frontier Max

Orchestration: directing agents is the skill

The build is bigger than an hour, or splitting work across agents is the thing being assessed. A bug hunt that fans several subagents out across the code, each one working to its own review brief, sits here.

- Multi-day take-home formats
- Live sessions built around orchestration
- Shows how somebody defines the contract between agents and reconciles what comes back

YOU GET THE BEST SIGNAL FROM INTERVIEWS LIKE THESE

01

A multi-file build with tests expected

Structure shows up here in a way it never does in a single-file answer.

02

A codebase-comprehension exercise

The candidate has to explain the system before touching it, so you see judgment before you see code.

03

A bug hunt against a real running service

Under time pressure you find out what somebody checks and what they assume.

ENTITLEMENTS REFLECT CURRENT PACKAGING. CONFIRM SPECIFICS WITH YOUR ACCOUNT TEAM

THE PACKAGE SETS THE CEILING ON WHAT YOU CAN SEE

Where you want to be

A team that sizes against the interviews they already run will land back on the package they already have. Size against the interview that reflects the job instead. For most engineering hiring that means a real build scenario, and the package decides whether a candidate is able to show you one.

WHAT AN INTERVIEW AT THIS LEVEL CAN REVEAL



LEVEL 0	LEVEL 1	LEVEL 2	LEVEL 3
STANDARD	FRONTIER LITE / FRONTIER	FRONTIER PRO – RECOMMENDED	FRONTIER MAX

LEVEL	PACKAGE	THE INTERVIEW IT UNLOCKS	WHAT THAT LETS YOU HIRE FOR
Level 0	Standard Starter · Team	Ask-only screens	Reasoning and judgment. Nothing about how somebody works inside a codebase.
Level 1	Frontier Lite • Frontier Team Plus · Enterprise	A single agent on a real repo	Engineers who can brief an agent and review what it returns. The floor for a coding interview.
Level 2 **RECOMMENDED**	**Frontier Pro** Package upgrade	**A full hour of real build work**	Anyone who will own work other people build on. Multi-file changes, tests, agents alone or in parallel, assessed on what they built.
Level 3	Frontier Max Package upgrade	Multi-hour and orchestration	Engineers who split work across agents and hold the pieces together. Extended formats, or orchestration as the assessed skill.

THE INTERVIEW IS THE CEILING

On an ask-only screen a strong engineer and an average one hand back the same thing, which is a conversation. Nothing got built, so there is nothing to compare. **The difference never gets a chance to appear.** A build interview makes the gap visible, and it needs a package that can carry one from start to finish.

FIND OUT WHICH PACKAGE YOU'RE ON

Your account team can map your current package to the interviews it carries, and show you what becomes testable at the next level.

coderpad.io