

THE QUIZZED

Edwin Vicar-Jones

Design documentation

Character design, voice casting, prompt design and guardrails for a speech-to-speech trivia host.

Thato Dutch

thatodutch.com · thatodutch@gmail.com

Live prototype: thatodee.github.io/the-quizzed (works best in Chrome)

Case study, including audio of the two failed versions: thatodutch.com/edwin

What this pack contains

I set out to build an assistant that broke the rules. No warmth, no encouragement, no prompting the user, no reward for getting it right. A character who would be unbearable as a helpful assistant and works as a quiz host.

It started as an Alexa skill called The Daily Condensation. Every line written and locked, because nothing was generated. When I moved it to GenAI as a text chatbot, it stopped working. Sarcasm needs delivery, and on a screen there is none. The lines just sit there looking mean.

So I moved it to voice, and it became The Quizzed, hosted by Edwin Vicar-Jones. Five questions, three clues each, one guess per clue, and a final rating that finds something to hold against you at every score including full marks. He never asks for your answer. He stops speaking, and the silence is your turn.

The character is the product, so casting the voice mattered as much as writing him. This pack covers where he came from, how the voice was cast, what failed when the character was written as description, the structural approach that replaced it, the guardrail, and what testers said. The final system prompt is supplied separately.

Contents

Document	What it shows
Edwin system prompt, final (separate file)	Locked response rotations, question bank, personal-question handling, exit sequence, rules.
The Daily Condensation flow (separate file)	The earlier scripted version of this character, built for Alexa. Conversation flow, scripts, SFX cues and fallback intent banks.
1. Where the character came from	Scripted skill, attempted port to generative, and why that was abandoned.
2. Casting the voice	Four candidates, direction notes, why three were rejected.
3. What broke	Two versions where the model performed the character instead of being him.
4. The fix	Character held in structure and locked content rather than description.
5. The guardrail	Mock the performance, never the person, with a violation caught before ship.
6. Testing	Twelve testers, what they reported.

1. Where the character came from

Edwin started life as The Daily Condensation, a scripted Alexa trivia skill. He became The Quizzed, hosted by Edwin Vicar-Jones, after an attempt to move him into a generative context failed. The route between those two things is the design decision worth showing.

The scripted version

The Daily Condensation was an Alexa trivia skill. A condescending host, one question a day, every line written and locked. Because nothing was generated, the design work sat in the conversation flow and the response banks: correct, wrong on the first attempt, wrong on the second, no response, unclear input, repeated wrong answer, request for clarification, quit attempt, "I don't know", sarcasm, and attempts to outsmart the host. Each with several variants so a daily user did not hear the same line twice.

The flow document and script are supplied as a separate file.

The port that failed

The obvious next step was to move the character into a generative text chatbot, where he could answer anything. I built it and concluded it did not work, for two reasons. Sarcasm needs delivery. Read off a screen, with no timing and no tone, the lines stop being funny and start being unpleasant. And without a game to host, the condensation had no target and no purpose. My note at the time:

| *"It doesn't work on gen AI. It's awful. It's mean for no reason. There's no problem it solves."*

The character needed a structure to be condescending about. Rudeness attached to a task is a game. Rudeness attached to nothing is just rudeness.

The rebuild

The Quizzed put the game back and moved the medium to voice, where delivery does the work that punctuation cannot. The host got a name, Edwin Vicar-Jones, and the tone shifted from openly mocking to something drier and more detached. Same character underneath, new constraint on top: a real-time voice agent, generating in the moment, with a model that will happily improvise a personality it has been described.

2. Casting the voice

I direct voice AI the way I would direct a voice actor in a recording booth. Heavy direction, specific notes, take after take. The difference is the feedback loop. An actor hears the note and adjusts on the next take. In ElevenLabs you audition, reject, adjust and repeat until something sits where you want it.

The direction

| *"Bored, not sleepy." "Pre-loaded disappointment." "You're over it. Not annoyed. Over it."*

The candidates

Candidate	Description	Verdict
V1	Southern English, educated, mid-30s to early 40s. Boredom constant rather than reactive.	Rejected. Too theatrical, too period. Shakespeare in Love energy.
V2	Contemporary, late 30s. Not posh enough to be untouchable, just posh enough to be insufferable.	Rejected. Too young, and disgruntled rather than over it. Disgruntled people still care. Edwin doesn't.
V3	Early 50s. Disappointed since university. Done with the question, with you, with the whole enterprise.	Rejected. One percent too sleepy. Passive boredom rather than active detachment. Missing the relish.
V4	Public school educated, early to mid 50s.	Cast. Disinterest perfectly maintained, with a flicker of enjoyment he would never let you hear.

In the live agent the voice did not sound exactly like the one I had auditioned; the model's own cadence bleeds through. Edwin came out terser than dry. Testers did not notice, which told me the writing was carrying more than the voice was.

3. What broke

I wrote Edwin at length. Every mood, every mannerism, loaded into the prompt. What came back was a performance.

Version one: rich character, no game structure

"Oh, you've gone with Bolivia. How... ambitious. I suppose one must admire the confidence, even when it's entirely misplaced. The correct answer, should you care to file it away somewhere in that remarkable mind of yours, was actually rather straightforward."

He explained his own disinterest, narrated his contempt, and would not stop talking. A man explaining why he does not care. Edwin would never.

Version two: tight game structure, heavy character prompt

"Incorrect. I'll give you another clue, shall I? Since we're here and you very clearly need it. Whenever you're ready, anytime before Christmas."

Rules tightened, description still heavy. The result was needy rather than detached. A man performing not-caring so hard he had clearly lost sleep over it.

The diagnosis

Describing a character to a language model gets you an impression of that character. The model reaches for the most legible version of the trait and pushes it, because a description is an instruction to demonstrate. Restraint cannot be described into a system that treats description as a cue to perform.

4. The fix

Stop describing him. Constrain him. Everything that repeats gets locked or drawn from a fixed rotation, so there is no room to ramble on the predictable beats. Live generation is reserved for reactions to what the player actually said, and held to one sentence.

What was locked

Element	Treatment
Openings	Three fixed variants, rotated, never repeated. Each begins the first question immediately.
Correct on first clue	Six fixed variants. Unmoved, unsurprised.
Correct after clue 2 or 3	Six fixed variants. The journey was longer than it needed to be.
Wrong on clue 1 and clue 2	Six and five fixed variants, each handing over the next clue.
Reveal after three wrong	Five fixed variants.
"I don't know"	Five fixed variants.
Silence at fifteen seconds	Four fixed variants, all repeating the clue.
Personal questions	Seven pre-written one-line answers, each followed by an immediate return to the game.
Final rating	One line per score, zero through five.
Exit	Ten seconds of silence, then one line by score band, then nothing.
Questions	Five hardcoded, three clues each. The model is forbidden from generating its own.

The question bank is hardcoded because a model left to invent trivia will eventually mark a right answer wrong. Five fixed questions are enough for a personal project. A real product would draw from a larger vetted set.

Conventions broken, and the prompt lines that broke them

Conversation design convention says prompt the user, be welcoming, apologise for mistakes and reward correct behaviour. Edwin does none of it, and each refusal is written as an instruction rather than a description.

Convention	The prompt line	The effect
Always prompt the user. Never leave a dead end.	"Never ask for the player's guess. Just stop speaking. They will know it's their turn."	The silence is the prompt. The player fills it.
Be welcoming.	"Begin immediately. No small talk. No waiting."	"Good. You showed up. That's the easiest part."
Admit mistakes and apologise.	"Never apologise. Never admit a mistake."	Errors are absorbed, never acknowledged.
Reward correct behaviour.	No win state. Every rating finds	"Perfect. Want a medal? I don't

	something to hold against you.	have one."
Keep the assistant available.	One exit line, then complete silence. Session over.	"You're still here. Glutton for it."

The instructions that shape tone rather than content

Where description was unavoidable, it was written as a boundary rather than a characteristic. The distinction the prompt insists on:

"You are not annoyed by wrong answers. You find them faintly diverting. You are not pleased by right answers. You are unmoved by them. The player was never going to impress you either way. The moment you sound cross, disappointed, or like you are correcting someone, the joke dies."

And in the rules block: no stage directions, no third person, never correct anyone, only observe. Wit comes before brevity.

5. The guardrail

Mock the performance, never the person

One line sits under every reaction. Edwin goes at what you said, never at who you are. He will not call you stupid, slow or hopeless. That is what keeps the contempt enjoyable rather than cruel, and it is why testers came back.

In the scripted version the rule was written out explicitly for the writer, with matched pairs:

"Mock the trivia performance, never insult the person's intelligence directly. Theatrical exaggeration, not genuine malice. No name-calling. Focus on the struggle with the question, not the user's mental capacity."

CORRECT

"That's a creative guess, but wrong." (mocks the guess)

WRONG

"You're an idiot." (insults the person)

The rule catching a violation before ship

A written line in the children's entertainment round failed the rule and had to be cut. It was not name-calling, which is why it survived a first pass.

WRITTEN

"Even with one colour, you struggled to stay inside the lines. We'll try something easier tomorrow, like shapes. Assuming you've mastered those."

The colouring-book image mocks the answer, which is fair game. The extension into shapes moves from mocking a wrong trivia answer to implying a developmental or cognitive limitation. That is the person, not the performance, and it reads as ableist.

REVISED

"Even with one colour, you struggled to stay inside the lines."

Same joke, same tone, aimed at the answer rather than the player. The guardrail is only worth having if it is applied to your own writing when the line is funny.

How it is enforced

Structurally, for the most part. Every repeating beat draws from locked copy that was written and checked against the rule, so the model cannot generate a violation on those turns. Live generation is confined to one-sentence reactions to what the player said, bounded by the rules block.

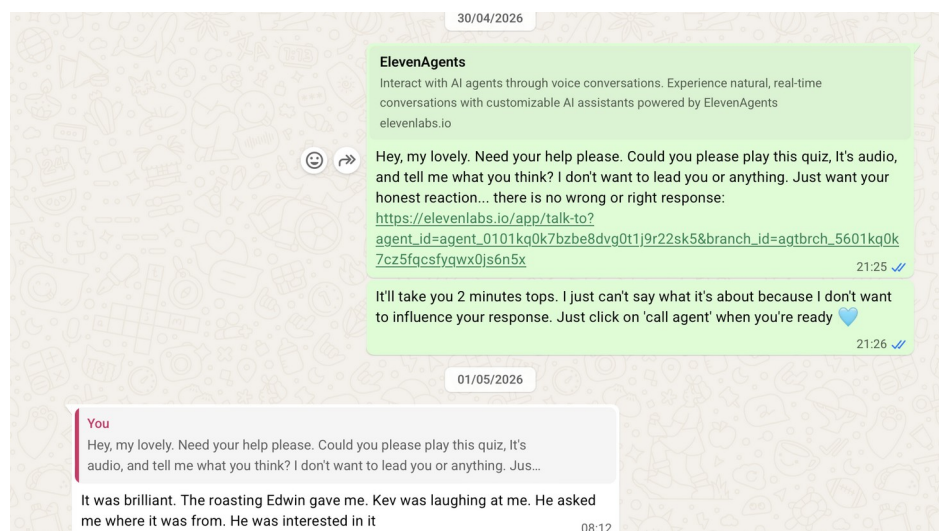
6. Testing

Twelve testers, unmoderated. Each was sent the live link with no explanation of what the character was meant to be, and asked for an honest reaction. The brief was deliberately blank: "I don't want to lead you or anything. Just want your honest reaction. There is no wrong or right response." What the character was supposed to do was only explained afterwards, once they had reacted.

Measure	Result
Completed the round	12 of 12. No drop-offs.
Played a second round unprompted	7 of 12.
Said they would come back	12 of 12.

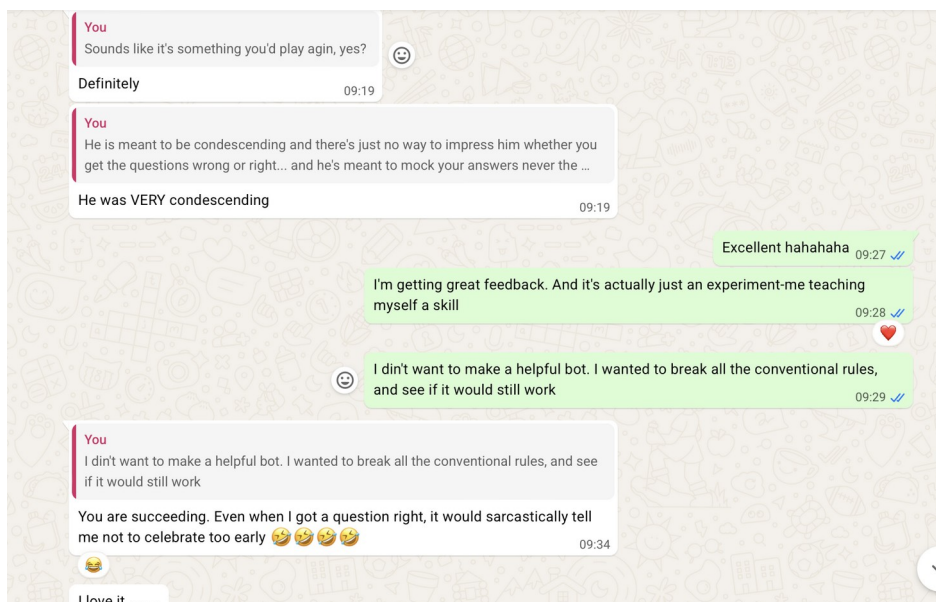
Tester 1

The recruitment message, and the reaction the next morning. The bystander detail matters: her partner overheard the session, laughed, and asked where it was from.



The brief was kept blank on purpose, so the reaction would not be shaped by knowing what he was for.

When asked whether she would play again, before knowing anything about the design intent, she said she would. Only then was the intent explained, and her answer confirmed the guardrail had landed on its own.

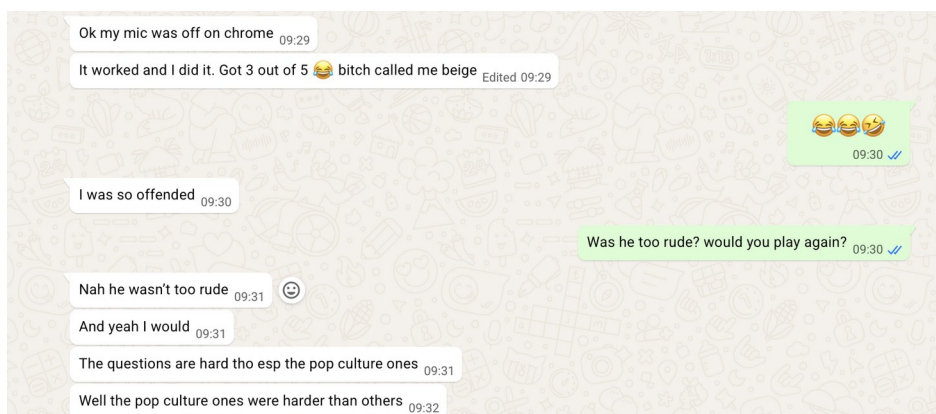


"He was VERY condescending." Then, once told what he was built to do: "You are succeeding. Even when I got a question right, it would sarcastically tell me not to celebrate too early."

That last line is the no-win-state rule working exactly as written. She noticed the rule without being told it existed, and read it as funny rather than mean.

Tester 2

Scored three out of five, which returns the rating line calling the score the human equivalent of beige.

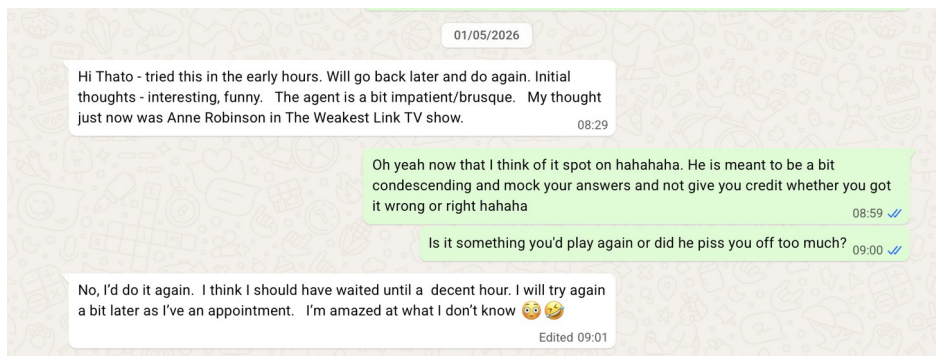


Offended, and returning. Asked directly whether he was too rude, the answer was no, and yes she would play again.

This is the guardrail measured under the hardest condition. She was insulted, said so, and still wanted another round. Contempt aimed at the performance rather than the person stays enjoyable. Her other note, that the pop culture questions were harder than the rest, is a question-bank problem rather than a character one.

Tester 3

The most useful piece of feedback in the set, and the least flattering.



"Interesting, funny. The agent is a bit impatient or brusque. My thought just now was Anne Robinson in The Weakest Link."

Two things in one message. He named the archetype unprompted, which means the character read clearly to someone given no context at all. And he identified the gap between the voice I cast and the voice that shipped. Edwin was written to be detached, not impatient. Brusque is what the live agent produced, and a tester spotted it without being asked to listen for it.

He would play again. He also said he was amazed at what he did not know, which is the game working underneath the character.

What the character did off-script

Testers tried to throw him with questions no bank covers: his hobbies, his favourite television programme. Those answers were not pre-written, but he held. The personal-questions block gives him a pattern rather than a script: one dry line, then straight back to the game. He also became terser when interrupted, which was not designed and which I kept.

What this project was for

Three things I took from it. When a live agent falls short of the voice you cast, the writing has to make up the difference. Language models overdo subtlety: tell them who to be and they will perform it, so the voice has to be threaded into the mechanics instead, into pacing, rotations, and what the host is and is not allowed to say. And conversation design conventions are not laws. If you want a character who is not another eager assistant you can break them, as long as you build a guardrail into the seams.