

Prompting Is Not a Design Skill. Writing Acceptance Criteria Is.

Getting better at asking does not help if you have no way to judge what comes back.

by Shreyas Kalyanpad

PUBLISHED: OCTOBER 19TH · BY SHREYAS KALYANPAD

Acceptance criteria = a written description of done that someone else could check without asking you what you meant.

Half the design internet is now saying that prompting is the new design skill. That take is one level too low. A prompt is a request. Criteria are a judgement. You can get better at asking forever and still not know whether what came back is any good.

The research points at the same gap. In the AI in Design Report 2026, 62 percent of designers name inconsistent or unreliable output as their biggest challenge. Reliable output is also the top reason a tool sticks, named by 80 percent. The thing designers want most is the thing they say they are not getting.

If you cannot write down what right looks like, you are not reviewing the output. You are reacting to it.



The checklist on the right is doing more work than the wall on the left.

1. Adjectives are not criteria

What we write: Make it clean and modern.

Why it fails: Clean has no failure state. Any screen can be defended as clean, including the one you hate.

Write instead: A limit and a number. Three type sizes. Body text at 16px or larger. Nothing under 4.5 to 1 contrast. Now the output either passes or it does not.

2. Describe the finish, not the request

What we write: A longer prompt, with more background about the brand.

Why it fails: Background shapes the attempt. It does not define the finish line. A report on spec driven development, published on arXiv in January, puts it plainly. These models are good at pattern completion and bad at mind reading.

Write instead: Two short lists, before you generate anything. What must be true. What must never appear.

3. I know it when I see it is the bug

What we say: Shali Nguyen leads consumer experience design at DoorDash. In that report she named the gap as fine grain control. Her words: I know what I want in my head, but you are just not giving me what I want.

Why it fails: That sounds like a tool problem. Usually it is an unwritten standard. The tool could not fail your test, because there was no test.

Write instead: Write the criterion the second you reject something. A rejection is free specification, and it is the one moment you actually know the rule.

4. Review is the job now

What we do: UX Tools surveyed 1,478 designers this spring. Only 1.4 percent trust AI output with no oversight. About a third trust it for production after a review pass.

Why it fails: Most of us review by feel, one screen at a time, against a standard that quietly drifts by Thursday.

Write instead: One checklist per surface. Same list, same order, every time. Boring on purpose.

5. Criteria outlive the tool

What we do: Collect prompt tricks for whichever model is winning this month.

Why it fails: The tricks expire with the model. The average designer now runs seven AI tools, up from three a year ago, so that is seven sets of tricks to maintain.

Write instead: Keep the criteria in your design system, not your prompt history. They move between tools unchanged. That same arXiv report cites early studies where human refined specs cut errors by up to half. It is honest that the evidence is thin so far.

Swap the adjective for the test

WHAT WE ASK FOR	WHAT WE SHOULD BE ABLE TO CHECK
Make it clean	Three type sizes, one accent colour, an 8px spacing grid
Make it feel premium	No stock photography, no gradient buttons, real product data
Keep it on brand	Colours only from the token file, no invented shades
Make it accessible	4.5 to 1 on body text, focus visible on every control
Make it consistent	Components from the library, no one-off variants

The audit

Open the last thing you generated and could not quite accept. Write the three sentences that would have told you it was wrong before you looked at it. If you cannot write them, the tool was never the problem. If you can, you have just written your first acceptance criteria, and you get to reuse them tomorrow.

Checked against

Read on 19 October 2026.

- AI in Design Report 2026, Designer Fund and Foundation Capital: 62 per cent of designers name inconsistent or unreliable output as their biggest challenge, 80 per cent say reliable output is what makes a tool stick, and the average toolstack has gone from 3 tools to 7. Includes Shali Nguyen of DoorDash on fine grain control
- UX Tools, State of Prototyping Spring 2026, 1,478 respondents: only 1.4 per cent trust AI output with no oversight, 32.8 per cent trust it for production with a review pass, and AI output quality is a blocker for 52.2 per cent
- Piskala, Spec-Driven Development: From Code to Contract in the Age of AI Coding Assistants, arXiv:2602.00180, 30 January 2026: models are good at pattern completion and poor at mind reading, good specs are testable and unambiguous, and early controlled studies show human-refined specs reducing errors by up to 50 per cent