Gemini's head of product: start evals at five prompts, and plan AI bets on a 60/20/20 split

PM Daily Digest

2026-11-09

Gemini's head of product: start evals at five prompts, and plan AI bets on a 60/20/20 split

By PM Daily Digest • November 9, 2026

Tulsee Doshi of Google lays out how AI product management is changing. Separately, operators argue that the line between PM and engineer is disappearing, and founders trade advice on validating an MVP that is already built.

Tulsee Doshi on what AI product managers actually do

On Lenny's Podcast, Google's Gemini product lead argued that when anyone can build easily, building something *great* is what sets you apart. Taste, meaning pulling signal from noise and finding the insight that unlocks an experience, is now the PM's core value [1]. Fast building also creates chaos, so the PM who drives consensus and brings structure matters more [1].

The practical points:

- **Hiring.** She wants PMs who use models every day, hold opinions, stress-test capabilities, dig into data and can build an eval. She also wants adaptability, because "a road map you wrote two months ago, you're going to be looking at it again and rewriting it" [1].
- **Start evals small.** Five prompts, or a mini-set of about 20, can show differences between two models. Act as the rater yourself to work out what "good" means before you scale up [1]. Evaluate end to end: judge quality, production filters and serving all change what users experience. In audio, for example, serving affects how natural the voice sounds [1].
- **Check that your metric represents your users.** Google's skin-tone evals once used the six-tone Fitzpatrick scale, which was built for healthcare and skewed light. Switching to Professor Monk's 10-tone scale changed how products were evaluated and fed into improvements in the

Pixel camera and image generation [1].

- **Model output is unpredictable.** Set capability goals, but the job is also to look at what comes out, find where the model shines unexpectedly, and turn that into something for users [1]. The model's behavior is UX too. Her team turns user feedback into a behavior spec, then into rubrics, evaluation data and reward modeling [1].
- **The 60/20/20 rule.** About 60% of what you build should work now. Another 20% should work well enough, with capabilities you expect to arrive soon. The last 20% is an ambitious target where you see "glimmers of light" [1]. Retest regularly: an older, rigid setup can hold back what a newer model can do [1].
- **AI-written docs.** Use them for data, such as post-launch usage and feedback roundups. The insight should come from the PM [1]. Her broader point is that the PM no longer just writes requirements for others to flesh out. The job now carries "some heft," whatever format that takes [1].

Fewer handoffs between understanding a problem and building the fix

A widely shared thread forecasts that the PM/engineer distinction will start to disappear. "PMs who can't build and engineers who don't understand customers will both struggle." Customer understanding becomes more valuable because "building 10 features is becoming easier than figuring out which 1 actually matters" [2]. Hiten Shah pointed to Monaco, where one of the best sellers now leads agentic deployment and the best GTM people help build the agents [3]. He says that after working with coding agents since February, he questions how much of a software org chart exists only because "the person closest to the problem couldn't build it" [3].

Shah also shared his team's setup, which looks a lot like a PM operating model: - One lead agent per project keeps the plan and records decisions. It delegates parallel tasks to other agents [4]. - Anything that needs human judgment goes into a single queue, and risky work needs explicit approval [4].

Validating an MVP that already exists

A founder with a launched MVP and no users drew many replies. The most useful one splits the problem into two questions. First, do people actually have the problem? Ask them to walk through the last time it happened. Second, does this version deliver value? Watch them try a task without coaching, then check a week later whether they came back or went back to their old workaround [5]. Another founder says his first 12 users came from DMs. He offered manual setup to frustrated people he found complaining on Reddit, and charged from day one, even $5: "If nobody will pay $5, that's your answer" [6]. Others said the signals to look for are an existing workaround [7] or money already spent on the problem [8].

Entry-level PM roles

In r/ProductManagement, one commenter said entry-level PM jobs have gone "from rare and super competitive to non-existent." He blamed overhiring in 2021–23 and AI agents doing entry-level PM work [9]. Another says junior BA roles, a common path into PM, are also becoming less common because of AI [10]. Both are personal views, not hiring data.

Also

Ryan Hoover says founders should introduce their company with "We are…" instead of "We're trying to…" [11].

Sources

1. Why PMs are suddenly valuable again | Tulsee Doshi (Head of Product, Gemini)
2. X post by @abishekv
3. X post by @hnshah
4. X post by @hnshah
5. r/startups comment by u/benjzammit01
6. r/startups comment by u/dmdbGroup
7. r/startups comment by u/Low_Cable2610
8. r/startups comment by u/AnonJian
9. r/ProductManagement comment by u/mikefut
10. r/ProductManagement comment by u/readyforgametime
11. X post by @rrhoover