

Shared Context, PM Eval Loops, and the New AI Fluency Bar

PM Daily Digest

2026-06-03

Shared Context, PM Eval Loops, and the New AI Fluency Bar

By PM Daily Digest • June 3, 2026

This brief covers the shift from 10x individuals to 10x teams, a practical eval-driven workflow for AI-assisted product work, and career signals around AI fluency and role design.

Big Ideas

- **The next AI advantage is team context, not just individual speed.** Ravi Mehta argues AI has created 10x individuals but also siloed “context engineering”; the next wave is 10x teams that turn shared context into shared progress [1]. He frames the canvas as the place where ideas, decisions, prototypes, and tradeoffs live together, and says the PM role is shifting from prioritizing what to build toward curating what should ship [1]. **Why it matters:** as artifact creation gets cheaper, alignment becomes the bottleneck. **Apply it:** keep decisions, prototypes, and tradeoffs in one shared workspace instead of splitting them across separate tools.

“The canvas is the team’s shared context.” [1]

- **AI-era PM leverage comes from defining quality faster.** Aparna Dhinakaran says top AI PMs stand out by defining what “good” looks like at scale faster than anyone else, and Aakash Gupta notes PMs who already run traces and evals on their agents are pulling away [2]. **Why it matters:** as code gets cheaper, the differentiator shifts to evaluation judgment [2]. **Apply it:** make eval design, failure analysis, and trace review part of normal PM practice for recurring AI-assisted work.

Tactical Playbook

1. **Start with one narrow PM agent, then harden it with evals.**
 - Build one job first: pull GitHub issues, score priority, and write a daily build report [2].
 - Inspect predictable errors early; in the example, bugs ranked too low while feature requests beat production issues [2].
 - Add even a noisy eval, then iterate: refine eval -> improve agent -> collect better traces -> sharpen eval [2].
 - Schedule the loop once it is useful; at Arize it now runs on cron [2].
Why it matters: triage compresses from weeks of manual review to fast feedback without removing PM judgment [2].
2. **For meeting recap automation, give AI the transcript and local context—not just notes.**
 - One PM found ChatGPT output generic because it did not know the existing PRD or what engineering expected [3].
 - A suggested fix was to record the meeting and feed the full transcript to AI along with context and explicit instructions for what to generate [4]. **Apply it:** include the current PRD, engineering expectations, and desired outputs in the same prompt [3, 4].

Case Studies & Lessons

- **Arize’s PM-agent demo shows where the new bottleneck sits.** In 45 minutes, one prompt processed 40 discussions, 60 issues, and 8 releases and scored them all [2]. The first output was not correct, but the errors surfaced immediately, were instrumented with evals, and improved fast enough that the loop now runs continuously and can support same-day issue identification, prototyping, and shipping [2]. **Key takeaway:** when build time collapses, the winning loop is *generate* -> *inspect* -> *evaluate* -> *iterate* [2].

Career Corner

- **AI fluency is becoming explicit.** Zapier’s open-sourced rubric scores four components—mindset, strategy, building, and accountability—across four levels from unacceptable to transformative, and it breaks expectations out by role, including product [5]. **Apply it:** use it for self-assessment, hiring, onboarding, or development planning. Full rubric [6]
- **Interrogate PM job design before you accept the title.** In one Reddit thread, a solo PM covering four products, no designers, and end-to-end work was measured on 90%+ on-time delivery, +20% sprint velocity, and NPS 8+ while being benchmarked as a project manager [7]. Commenters said velocity and milestone targets are project-management metrics, called the KPI mix contradictory or gameable, and urged job hunting; some also described \$85k as far below the scope, citing higher ranges from their own

experience [8, 9, 10, 11, 12]. **Apply it:** ask whether the role truly owns strategy and outcomes—or mostly backlog, deadlines, and coordination.

Tools & Resources

- **Miro Canvas 26** is worth tracking if your team struggles with fragmented context. The product direction centers on Sidekicks that pull team context, Flows that turn messy workshop inputs into structured outputs while keeping humans in the loop, live Code to Prototype on the board, and MCP to make the canvas readable and writable by AI agents [1].
- **ProductHQ** is a new PM-built workspace aimed at reducing tool sprawl across discovery, prioritization, roadmapping, PRDs, spec generation, and Jira push, with an optional AI assistant called Maya. It is free to try at myproducthq.com [13].

Sources

1. From 10x people to 10x teams
2. substack
3. [r/prodmgmt](#) post by u/Independent_Bee_9219
4. [r/prodmgmt](#) comment by u/greenbeen
5. X post by @sachinrekhi
6. X post by @sachinrekhi
7. [r/ProductManagement](#) post by u/Organic_Ad9083
8. [r/ProductManagement](#) comment by u/utzutzutzpro
9. [r/ProductManagement](#) comment by u/cderm
10. [r/ProductManagement](#) comment by u/theecstaticdawning
11. [r/ProductManagement](#) comment by u/ratczar
12. [r/ProductManagement](#) comment by u/londongastronaut
13. [r/prodmgmt](#) post by u/Longjumping_Dog_883