

Munish Raizada, Vice President, Primus Partners

Published in Financial Express
Mar 19, 2025 | 10:08 AM IST

Meta's Community Notes: A tool for transparency or a gateway to chaos?



Authored by Geetika Srivastava

Read on: <https://www.financialexpress.com/business/brandwagon-metas-community-notes-a-tool-for-transparency-or-a-gateway-to-chaos-3781483/>

Article Content:

In India, where misinformation spreads rapidly across social media and messaging apps, relying on a crowdsourced model for fact-checking is a double-edged sword. Without an independent verification process, there's no guarantee that Community Notes will always improve content accuracy.

For brands, this creates a new challenge. A marketing campaign could be flagged as misleading simply because of misinterpretation, regional language barriers, or conflicting opinions. Unlike independent fact-checking, which follows structured research methods, Community Notes could amplify subjective viewpoints instead of facts.

It has the potential to lead to chaos, especially in India, where online narratives can influence real-world behaviour. The key challenge is to prevent the system from being gamed by groups looking to manipulate narratives.

For businesses and advertisers, this adds unpredictability. Content moderation follows a clear process, but now brands may have to navigate public perception in real time. A campaign that is factually correct but misinterpreted by a large enough group could face unnecessary scrutiny.

Meta's earlier model, which worked with third-party fact-checkers, at least followed a structured approach based on verified sources. Replacing this with a user-driven model introduces the risk of content being judged based on perception rather than accuracy.

For brands, this means an added layer of complexity. Earlier, if content was flagged, there was a clear process for appeal. Now, brands may need to constantly monitor and counteract misinformation through direct engagement, which can be time-consuming and difficult at scale.

Since Community Notes has worked reasonably well on X, there's a possibility it can be effective on Meta's platforms too. But India's digital landscape is more complex, requiring additional checks to ensure accuracy across multiple languages and regions. For brands, advertisers, and content creators, this shift means staying more proactive in managing narratives and ensuring their messaging isn't misinterpreted or flagged unfairly. Independent fact-checking remains crucial in a market as diverse and dynamic as India, and Community Notes should ideally complement, not replace it.

Another layer for CMOs to navigate

Ambika Sharma, Founder & Chief Strategist, Pulp Strategy

Meta's latest move—introducing Community Notes — aims to tackle misinformation through crowdsourced fact-checking. In theory, this feels like a step towards greater transparency. But are we truly moving towards a more informed social space, or just opening the floodgates to more noise?

Right now, I see more potential for chaos than clarity. Without proper moderation and safeguards, brands, CMOs, and content strategists will have to navigate yet another unpredictable layer of digital discourse—one where “truth” is dictated not by facts, but by whoever can game the system best.

And let's be clear—it's not just about users getting things wrong. We need to account for amplification farms, bots, and agenda-driven groups. Anyone who has worked in digital strategy knows how easily online discourse can be manipulated. If Meta doesn't build strong safeguards, we could see misinformation countered with more misinformation, wrapped in the illusion of credibility.

I can't help but compare this to YouTube's content moderation system, where flagged videos undergo internal review before action is taken. While that process is slower, it at least ensures a level of oversight. In contrast, Meta's approach feels like an open forum where the loudest voices—not necessarily the most factual ones—win.

If Community Notes are to work, Meta needs to filter contributors carefully—not just based on engagement, but actual domain expertise. Moreover, they must create safeguards against bot networks and agenda-driven brigading that could manipulate fact-checking.

They should also combine human input with AI-backed verification to ensure real accuracy, not just perceived credibility.

Opens doors to bias and manipulation

Munish Raizada, Vice-President, Primus Partners

The need for a better approach to misinformation cannot be denied. Multiple studies have shown that misleading content spreads much faster than factual information. Meta has previously tried third-party fact-checkers like Reuters and AFP, but these initiatives were limited in reach and often slow. Transparency reports reveal that fact-checked content made up less than 1% of all flagged posts. This means that a vast amount of misinformation slipped through. Community Notes offers a more scalable solution that will allow everyday users to step in and provide additional context to potentially misleading posts.

Reports suggest that Community Notes, when tested on X, reduced engagement with flagged misinformation by nearly 40%. If Meta also manages to implement the system effectively, it could empower users to challenge false claims in real-time. Though this sounds interesting, the real challenge will lie in the execution. Although, crowdsourced

fact-checking allows diverse perspectives, it also opens the door to bias, manipulation and even misinformation being disguised as “corrections”.

Another critical concern that arises is whether people will actually engage with the system, as it is being expected. We have already seen on X, the participation rates in Community Notes have been quite low. This raises concerns about effectiveness and efficiency of this tool on Meta's platforms also.

Moreover, if Community Notes becomes a battleground where different groups push competing narratives, the feature might do more harm than good, eroding trust in fact-checking rather than strengthening it. There should be a transparent process to identify who can contribute to community notes.