Meta Response to Board Recommendations (Advice on Pharmaceutical Drugs Case)
2021-015-FB-UA

April 1, 2022

Recommendation 1 (implementing fully)

Meta should publish its internal definitions for "non-medical drugs" and "pharmaceutical drugs" in the Facebook Community Standard on Restricted Goods and Services. The published definitions should: (a) make clear that certain substances may fall under either "non-medical drugs" or "pharmaceutical drugs" and (b) explain the circumstances under which a substance would fall into each of these categories. The Board will consider this recommendation implemented when these changes are made in the Community Standard.

Our commitment: We will publish definitions for both non-medical drugs and pharmaceutical drugs to clarify how we apply our policies to these two types of restricted substances.

Considerations: We provide extensive, detailed resources to our content moderators to help them consistently apply our policies. The detailed guidance includes information such as examples, definitions and criteria relating to what kinds of content are allowed. We are working to publish more of these details so our enforcement protocols are more transparent. As part of that effort, we will make the definitions of "non-medical drugs" and "pharmaceutical drugs" in our Restricted Goods and Services policy publicly available.

These definitions clarify the existing policy distinctions between non-medical drugs and pharmaceutical drugs.

Next steps: We anticipate publishing these definitions in our Transparency Center this spring.

<u>Recommendation 2 (assessing feasibility)</u>

Meta should study the consequences and trade-offs of implementing a dynamic prioritization system that orders appeals for human review, and consider whether the fact that an enforcement decision resulted in an account restriction should be a criterion within this system. The Board will consider this recommendation implemented when Meta shares the results of these investigations with the Board and in its quarterly Board transparency report.

Our commitment: We will explore a dynamic prioritization system for our human review of appeals. As part of this effort, we will consider whether the fact that an enforcement decision resulted in an account restriction should be a criterion for prioritization.

Considerations: We generally review appeals in the order we receive them, though we do review people's appeals for their content we've removed for violating our policies ahead of appeals of decisions when they report other people's content.

We started considering new methods of prioritizing appeals when the COVID-19 pandemic introduced capacity constraints that required us to pause our plans. In line with our response to recommendation #2 in the Punjabi concern over the RSS in India case, we have resumed planning to assess dynamic prioritization models for appeals review.

As part of that work, we will consider whether to use account restrictions as a factor in prioritizing appeals. On the one hand, it seems reasonable to prioritize appeals when the penalty is the most restrictive — namely, excluding someone from accessing the platform. On the other hand, because account restrictions often result from repeated violations, prioritizing those appeals

could mean giving prioritized review to people who are most frequently violating our policies.

Next steps: We will start to assess the dynamic prioritization of appeals as part of our roadmap planning for the second half of 2022. We anticipate providing more information in the next Quarterly Update.

Recommendation 3 (implementing in part)

Meta should conduct regular assessments on reviewer accuracy rates focused on the Restricted Goods and Services policy. The Board will consider this recommendation implemented when Meta shares the results of these assessments with the Board, including how these results will inform improvements to enforcement operations and policy development, and summarize the results in its quarterly Board transparency reports. Meta may consider if these assessments should be extended to reviewer accuracy rates under other Community Standards.

Our commitment: We already collect and assess data on the basis of takedowns and restorations — including takedowns under our Restricted Goods and Services policy — to develop our classifiers, review protocols and policies.

As part of our commitment to transparency, we will continue to explore ways of adding appropriate accuracy metrics to the Community Standards Enforcement Report. Based on external stakeholder input and regular feedback (including that of the board), we believe that consistent expansion of the amount and type of data we publish is the best way to provide meaningful insight into our policies and enforcement efforts.

Considerations: As noted above, we collect data on the basis of takedowns and restorations to feed into continuous development of our classifiers, review protocols and policies. We assess success in terms of both reviewer accuracy

and policy clarity and scope. If we identify performance issues in the course of these continuous assessments, we attempt to identify the source of the issues. Depending on the source of the issue, we will involve the appropriate team to address it. These feedback loops may result in changes to our policy or enforcement protocols. For example, this may lead to updated or additional training for content reviewers.

We publish information concerning the prevalence of various types of policy violations, and our actions in removing or restoring content, in our Community Standards Enforcement Report. We will continue to pursue this strategy of publishing comprehensive information rather than prioritizing narrow segments of data that do not provide a complete picture of our content moderation successes and challenges. In line with previous board recommendations, we have expanded and accelerated our efforts to share enforcement accuracy rates. Our long term goal of broad transparency remains unchanged.

Next steps: We'll continue to assess transparency measures for reporting consistent, comprehensive and accurate data. We'll provide additional information as we complete these ongoing assessments, including as they relate to our Restricted Goods and Services Policy.