

Ethical Obligations to Provide Novelty

Paige Golden

Philosophy, Politics, & Economics
University of Pennsylvania
Philadelphia PA USA
email@email.com

David Danks

Philosophy and Psychology
Carnegie Mellon University
Pittsburgh PA USA
ddanks@cmu.edu

ABSTRACT

TikTok is a popular platform that enables users to see tailored content feeds, particularly short videos with novel content. In recent years, TikTok has been criticized at times for presenting users with overly homogenous feeds, thereby reducing the diversity of content with which each user engages. In this paper, we consider whether TikTok has an ethical obligation to employ a novelty bias in its content recommendation engine. We explicate the principal morally relevant values and interests of key stakeholders, and observe that key empirical questions must be answered before a precise recommendation can be provided. We argue that TikTok's own values and interests mean that its actions should be largely driven by the values and interests of its users and creators. Unlike some other content platforms, TikTok's ethical obligations are not at odds with the values of its users, and so whether it is obligated to include a novelty bias depends on what will actually advance its users' interests.

CCS CONCEPTS

- Information systems applications---Collaborative and social computing systems and tools---Social networking sites; 500
- Collaborative and social computing systems and tools---Web searching and information discovery---Content ranking; 500

KEYWORDS

Recommender systems, TikTok, Ethical obligations, Interest-based analysis

ACM Reference format:

Paige Golden and David Danks. 2021. Ethical Obligations to Provide Novelty. In *Proceedings of 2021 ACM Conference on AI, Ethics, and Society (AIES '21)*. May 19-21, 2021, Virtual Event. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/3461702.3462555>

1 Discovering Content

One impact of the rise of online social networks has been a decentralization of content creation. People are increasingly able to consume and respond to content that was produced by

individuals who are not part of the traditional content creation and dissemination systems. On the one hand, this decentralization has made it possible for more diverse voices to contribute content to the public sphere. On the other hand, it has led to a massive increase in the sheer amount of content, thereby potentially making it harder to find new or different voices.

As a result, essentially all content platforms employ algorithms (of varying sophistication) to determine which content is shown to particular users. Ethical questions naturally arise about the obligations (if any) on the content platform with respect to these recommender algorithms. For example, there have been a number of recent arguments about whether Facebook has any ethical obligation to minimize the likelihood of "information bubbles" [2]. And of course, numerous recent events have shown the real-world impacts of algorithmic individualization and curation of content to produce increased polarization and extremism. These are not purely abstract, theoretical questions, but rather connect directly with tangible harms and benefits for many different groups and individuals, including ones not even on that platform.

In this paper, we focus on a newer content platform that has not received much ethical scrutiny: TikTok (though see [1]). TikTok is unlike many other content platforms in hosting relatively little content intended to directly convey information, points of view, or advertising (though there was some use for political information and messaging during the 2020 election; see [7]). For many people, part of the appeal of TikTok is precisely that the content is relatively "light": users can quickly consume content and then proceed with other aspects of their lives.

Like other content platforms, TikTok uses recommendation algorithms to provide a For You feed that offers content that has been tailored (in some sense) to the user. Obviously, the company has practical incentives to include (in the For You feed) content that it predicts the user will like or enjoy. We ask a complementary ethical question: Does TikTok have any moral obligation to show diverse or novel content that might not rank as highly in terms of predicted enjoyment?

On the one hand, one might argue that TikTok has no moral obligation to do anything other than show the content that is predicted to be most enjoyable. This argument is naturally grounded in the idea that TikTok has no ethical obligations beyond those that arise from its corporate mission. On the other hand, as with concerns about Facebook and information bubbles, one might argue that TikTok should use its platform to provide users with exposure to diverse content and creators, thereby (perhaps) increasing tolerance and understanding of others. These obligations naturally arise if one thinks that corporations bear

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

AIES '21, May 19–21, 2021, Virtual Event, USA
© 2021 Copyright is held by the owner/author(s).
ACM ISBN 978-1-4503-8473-5/21/05.
<https://doi.org/10.1145/3461702.3462555>

ethical obligations beyond those that arise from their corporate values and mission.

We provide a different analysis and argument that does not attempt to adjudicate whether TikTok has these kinds of additional ethical obligations to advance social good. Rather, we contend that TikTok’s own stated values imply that it should support its users and creators, and so the ethical implications of a potential novelty bias in TikTok’s For You feed depend on open empirical questions. More specifically, we focus on the competing and complementary values and interests of key stakeholders around TikTok. We defend our interest-based approach (rather than a rights-based approach) in Section 2, and show how it raises key questions about psychological impacts of a novelty bias on TikTok users and creators. We then consider the values and interests of each stakeholder group separately in Section 3. Our approach in Sections 2 and 3 leaves open difficult questions of how to weigh and balance the competing interests with different moral weights; we return to this problem in Section 4 to offer an “all things considered” recommendation.

2 Values, Novelty, & TikTok

One approach for an ethical analysis would consider the relevant rights, duties, and responsibilities of the individuals and groups that use, create, consume, and spread content on TikTok, as well as connected corporations and institutions. However, as with so many issues involving novel technologies, there is little-to-no consensus about exactly what those rights, responsibilities, and duties might be (beyond high-level ones such as basic human rights that are not helpful in this context). We thus instead focus on the morally defensible values and interests of these groups (and so potential rights and/or obligations), with the recognition that these might vary between members of a stakeholder community.

When thinking about the stakeholder values that are potentially impacted by content that is created and consumed, it is important to distinguish between the content having intrinsic versus instrumental value. The intrinsic value of an object or event is the entity’s worth in and of itself. For example, basic human rights are often understood to have intrinsic value rather than being important merely as a means to an end. Alternately, an entity has instrumental value when its worth is due to its ability to help one reach a secondary goal. Most money is a good example of something with purely instrumental value: the physical paper has very little value in itself, but money can be used to purchase goods and services, signal social status, and/or live comfortably. And of course, the value of an entity or experience can be quite complex, encompassing both intrinsic and instrumental value. An example is a college experience: while learning certainly has significant value in and of itself, a degree from a university also has critical instrumental value for obtaining a job.

The TikTok For You feeds potentially provide users, creators, and the company each with both intrinsic and instrumental value from novelty (as we explore in Section 3), though the nature of the value differs from party to party. Full understanding of those values first requires an explanation of the particular details of the

TikTok For You feed. In general, recommendation systems (including content recommenders) aim to determine which content is most likely to be enjoyed by the user. Essentially all recommender systems will present content that the user has not previously seen, but such content will typically be quite similar to content that the user previously saw and enjoyed, as that content is most likely to receive a high predicted probability of enjoyment. That is, basic recommender systems typically produce content streams that are relatively uniform (along relevant dimensions).

A *novelty bias* in a recommendation algorithm increases the viewing probability of content that is unlike (in some sense) the content that a given user has either previously seen, or has expressed interest in. That is, a novelty bias—by definition—causes users to see content that a basic recommender system would predict they are *less* likely to enjoy (compared to other content that could potentially be shown to them). Moreover, a novelty bias will be strongly correlated with increased diversity in the content feed, as it will up-weight content that is different from what was previously seen (and enjoyed). Although such a bias is only a small part of the larger algorithm, and might only cause novel content to show up occasionally, it has become a recent topic of debate.

The precise design and implementation of the TikTok For You recommender system has not been publicly disclosed. However, observed patterns strongly suggest that the feed is strongly shaped by similarity, along multiple dimensions, to previously viewed content. There is some reason to think that similarity is measured in ways that can readily lead the feed to focus (for a user) on content created by individuals with similar demographic characteristics [3]. That is, the For You feed appears to behave similarly to many other recommender systems, and likely with a limited (at best) novelty bias. Does the company have a moral obligation to increase (or introduce) a novelty bias into the For You feed (thereby almost certainly increasing the diversity of feed content and observed creators)? The answer partly depends, we argue, on two open empirical questions.

The first empirical question concerns *user* response to content that is presented due to an algorithmic novelty bias (i.e., content that was not top-ranked in probability of user enjoyment): Do users gain value from this content because of its novelty? Currently, there is no consensus in the field with regards to the extent that users value novelty. A study addressing the dynamic nature of user novelty preferences found significant between-user variance in preference for a novelty bias in a basic recommender system; the study’s conclusion was that algorithms should thus employ user-specific novelty biases [4]. However, this study leaves open the question of value from that novelty.

Importantly, this empirical question does not require us to distinguish whether the value (if any) of novel content is intrinsic rather than instrumental. The relevance to TikTok’s moral obligations (if any) depends solely on whether there is *some* additional value for content if it is selected due to a novelty bias, no matter the nature of this value. In particular, is the value of novel content consistently greater than the expected value of other content for this user? Unfortunately, this question is very difficult to answer empirically. One possibility would be an A/B-type

experiment in which TikTok presents different users with For You feeds constructed with different degrees of novelty bias (from complete to absent). Alternately, TikTok could (for different users) adjust over time the degree of novelty bias within a user's feed. The social engagement of users could be monitored, perhaps along with explicit value judgments: "how much did you enjoy the content presented to you?" The recorded social engagements and user answers would be informative about the value of unique, novel content, though this experiment would likely require massive amounts of data. To the best of our knowledge, no such study has been conducted (or at least, no such study has been publicly discussed or published).

The second empirical question concerns *creator* response to novelty and diversity: Do creators receive additional value if the recommender system employs an algorithmic novelty bias? As with the open question about users, we do not know whether creators derive additional value from a novelty bias in content feeds. On the positive side for creators, a novelty bias should increase the chance that additional users (beyond the core audience for whom the predicted probability of enjoyment is highest) will see the content. On the negative side, the novelty bias can reduce the chance that their core audience will encounter their content. The net value from this tradeoff for creators is unknown, and probably depends on key algorithmic and platform details. Moreover, the net value will likely depend on the answer to the first key empirical question, since the value-for-users of novel content will influence whether they provide increased likes, views, and shares.

These empirical questions provide critical constraints on our ethical analysis because answers would help us to better understand and gauge the nature of potential harms from inclusion or exclusion of a novelty bias. In this context, we can distinguish between harms of commission and harms of omission. Harms of commission are those that occur as a result of a specific action. For example, a change from a flat tax rate to a progressive one would (likely) cause a harm of commission to wealthy individuals, as they would have to pay more taxes. In contrast, harms of omission occur when the stakeholder is denied access to a better alternative. For example, a government *not* changing from a progressive tax rate to a flat one would potentially be a harm of omission for wealthy individuals. The failure to change the tax code does not directly make them worse off than in their current state, but they are not benefited as much as they could be.

As these examples make clear, harms of commission and omission (and whether they are harms at all) depend on the relevant moral baseline for that individual or group. One only experiences a harm of omission, for instance, when denied access to a better alternative to which they are morally entitled or permitted. For example, one might argue that there is a harm of omission from the government's failure to provide all citizens with a private jet, but there is clearly no such (moral) entitlement, and so no harm from the government's failure to provide private jets. Although this example is clearly fanciful, baselines of moral entitlement can be quite tricky to establish for more complex issues. We revisit this question of the appropriate moral baselines

for different stakeholders in the context of novelty in app feeds in Section 4 of this paper.

3 Values, Interests, & Preferences

We now turn to the core values and interests of the three different core stakeholder groups: users, creators, and the company.

3.1 Users

TikTok users are clearly a highly heterogeneous group with diverse interests and values. Nonetheless, for the purposes of the present analysis, we treat them as a single stakeholder group since we are operating at a relatively high level of description, with the understanding that there are likely to be exceptions for each of our claims.

When we posed the first empirical question (i.e., about value for users), we remained agnostic about whether the potential greater value of novelty for users derives from intrinsic versus instrumental factors. This greater value might be due to the user intrinsically valuing novelty for novelty's sake, or it might be due to instrumental values such as the novel content introducing the user to types of content that they had not previously encountered (and that they liked more than the algorithm predicted). The greater value for the user could even be due to a longer causal chain, such as the novel exposure leading them to become more adventurous in their content consumption, thereby encountering other enjoyable types of content.

When we turn to analyzing the (morally relevant) values and interests of users as a stakeholder group, however, we can no longer remain agnostic, but instead must distinguish between different types (or reasons) of potential value for users. Novelty in a feed would have intrinsic value for a user to the extent that the individual enjoys new content purely for its novelty or differences from previously enjoyed content. That is, some users are presumably interested in scrolling through a feed full of unique content that is unlike anything they have previously seen.

A novelty bias also likely provides instrumental value for some users. Tailored content in the For You feed often features creators from the same geographical area as the user. In contrast, novel content is more likely to stray from the individual's location, exposing the user to other areas, cultures, and lifestyles. Exposure to other people's lifestyles allows for a greater world understanding and empathy for others' situations, which can be instrumentally valuable for some users. Increased novelty can also provide instrumental value through users' exposure to, and pursuit of, potential new interests and hobbies through novel For You page content. If users are unaware of various activities, then they clearly cannot pursue them. Exposure to new interests gives users the opportunity to discover hobbies that they may like even more than their current ones, leading to increased enjoyment.

Of course, presentation of novel content can reduce the likelihood that users see content that is highly pleasurable for them. Many TikTok users greatly enjoy perusing a feed full of content tailored to their interests based on their previous likes and shares (hence, their willingness to spend substantial time on

TikTok). Novel content in the For You feed can potentially be unpleasant, or even offensive, for some users (i.e., cause a harm of commission). There are risks of harm for people when showing them content that the recommendation system predicts is less likely to be enjoyed.

At the same time, this possibility must be balanced against potential harms of omission from *not* implementing a novelty bias. If users value novelty (i.e., if the first empirical question is answered in the affirmative), then they will be harmed by omission if no change is made, as there will be a lack of novel content in their feeds. (We momentarily suppose that they have a reasonable moral entitlement to content that they value, but revisit this issue below.) Similarly, the potential indirect benefits of novelty exposure—for example, the opportunity to gain a more worldly understanding through exposure to more diverse content—would be missed by a failure to implement novelty bias, potentially causing further harms of omission. Additionally, continuing to allow the For You page to lack novelty indirectly harms the user by depriving them of their ability to use TikTok to find and engage in new interests.

These observations all raise questions about the proper baseline of moral entitlement, which themselves require answers to two key questions: Are users morally entitled to *maximal* satisfaction from novelty in TikTok? And if so, then does TikTok know how to provide it to them? We consider the first question in Section 4; we conjecture that the current answer to the second question is negative. Based on public research and statements, we conjecture that TikTok is unaware of how to provide users with maximal satisfaction. If we are correct in our conjecture, then TikTok finds itself in a morally tricky situation: they lack the knowledge for how to best act (from the perspective of their users), but they could plausibly remedy that lack of knowledge through experimentation and testing. If our conjecture is correct, then TikTok should immediately take steps to better understand the intrinsic and instrumental values (if any) of novelty for their users, as well as which types of novelty are most valuable.

3.2 Creators

As noted earlier, there is a key open empirical question: What value do creators receive from a novelty bias in TikTok? The main interest of public TikTok creators is gaining likes, views, and shares. These different elements can lead to different kinds of value depending on the creator's interests and preferences. Once again, we must distinguish between intrinsic and instrumental values. Some creators find intrinsic value in having people view and spread their content. All benefits of fame aside, people who intrinsically value views simply enjoy having people see what they have produced and who they are, and the intrinsic value of likes and shares comes from the pleasure of one's content receiving approval from others.

Many creators also receive instrumental value from increased social engagement with their content. One of the main objectives of increasing engagement is personal fame and attention. Higher numbers of views on a video indicate one's status as a creator and potential as a wide-scale influencer. This is different from the

intrinsic value of views in that it is more focused on the admiration of other people than personal enjoyment in having your work seen. In addition, likes, shares, and views can enable users to promote messages that are important to them. For example, a creator put out a TikTok video accusing Wayfair of child trafficking, since this creator was trying to promote a petition and an important message (which required video engagement). A final instrumental motivation for accumulating views is to make money. While estimates on the exact amount that TikTokers are paid per million views greatly varies, their income rate is certainly dependent on their viewership. Since social engagement indicates that people are viewing and enjoying a creator's content, companies are more likely to sponsor creators if they have more likes, views, and shares.

Although we understand the proximal causes of value for TikTok creators, the open empirical question is exactly how an algorithmic novelty bias impacts these different causes: that is, do creators experience a net increase in value (for any or all of the above reasons) when a novelty bias is used in a content recommender system? Obviously, since creators derive some value from users viewing and liking their content, their value given a novelty bias is partly dependent on the value of a novelty bias to the users (i.e., the first open empirical question). If users are more likely to enjoy the novel content that is put in their feed and engage with it in the form of views and likes, then creators will have their values fulfilled by novelty in the For You feed.

However, due to the uncertainty of the answer to the empirical question about users, no firm conclusion can be drawn with regards to the values of novelty for creators. In terms of harm, creators could experience harms of commission in the form of decreased likes and views if users do not actually value novelty in the For You feed, and so engage less when a novelty bias is used. Harms of omission will occur if the user *does* value novel content but TikTok fails to use a novelty bias, as creators will miss out on the increased likes, views, and shares that come along with increased novelty. In short, the open empirical question about users' value from novelty is critical in assessing whether and when creators experience harm from TikTok's For You algorithm. The second empirical question (about values for creators) largely reduces to the first one (about values for users).

3.3 TikTok

Finally, we must consider the case of TikTok as a company. TikTok is an enterprise composed of many individuals, and so we focus simply on values for this collective, regardless of whether they are intrinsic (for individual members) or instrumental (for the group). In general, the main interest of any profitable content platform company is to gain and retain customers or users. Consequently, the values of a company tend to greatly depend on the values of its clientele. In the case of TikTok, their main interest is presumably to please the users and creators to get them to keep using and downloading the app. Despite this general core interest, the minutiae of the relationship between TikTok and its users and creators must be examined to get a more detailed understanding of the company's specific interests.

One way to begin examining TikTok’s values is in its mission statement which outlines the values that the company claims to have and pursue. TikTok officially states that their mission is to “inspire creativity and bring joy.” This mission statement suggests that even if novelty fails to fulfill any of TikTok’s profit- or engagement-centric interests, a novelty bias will be valuable for TikTok if it inspires creativity in creators, or brings joy to users. Of course, as the key open empirical questions note, we do not know whether users do derive value (“joy”) from novelty, nor whether creators are “inspired” by a novelty bias. TikTok as a company needs to know the answer to these difficult empirical questions in order to satisfy its own mission statement. And if the answers to those questions are affirmative, then the app’s mission statement suggests that TikTok must also value a novelty bias.

Some of TikTok’s other main interests are: attracting new users and creators, retaining current users and creators, and avoiding bad press. The potential value of a novelty bias thus lies in two, not mutually exclusive, possibilities: (a) increased user engagement—both attraction and retention—if the novel content is well received, and (b) reduced likelihood of negative press due to high-profile examples of For You feeds that fail to be inclusive towards other cultures, races, lifestyles, and ideas.

First, as with the instrumental value of novelty to creators, TikTok’s value of novelty is dependent on the intrinsic value of novelty to users. If users enjoy seeing novel content, engagement will increase. This increased engagement is a means to satisfy TikTok’s interest in bringing in new app users and retaining their previous ones. Increased engagement indicates that users are finding more enjoyment in their For You page, which helps TikTok’s interest in keeping users on their platform. A similar effect can be observed with regards to creators, who are also more likely to continue using the platform if it fulfills their interests of getting likes, shares, and views. With regards to bringing in new users, the number of TikTok downloads are also more likely to increase with increased user satisfaction. Overall, increased engagement due to users’ intrinsic value in novelty will keep users and creators interested in the app and bring in new app users.

However, if users do not enjoy seeing novel content on their feed, they will show decreased engagement with the videos. In this case, TikTok will experience both harms of commission and omission. With regards to the former, if user engagement decreases after novelty is implemented, then users will likely become bored with their feeds, causing them to like and view less content. Consequently, creators will feel frustrated when their likes and views decrease. This will result in TikTok losing app users and receiving fewer new downloads due to poor user reviews. With regards to the latter, a harm of omission will occur if users intrinsically value novelty but TikTok fails to implement an increase in it. Such a failure to act will result in a significant missed opportunity to elevate the user experience and receive more downloads. If users do value novelty in and of itself, TikTok is missing out on the opportunity to use novelty to increase engagement.

TikTok’s second interest with regards to novelty is its likely impact in reducing negative press coverage. TikTok has recently received substantial criticism with regards to race issues [5].

Whether these accusations are based in fact is relatively unclear due to the lack of hard (public) data on the matter. Either way, the opinion of many users tends toward the idea that TikTok’s lack of representation of many different races and cultures in the For You feed indicates a racist trend. In fact, TikTok itself released a formal statement apologizing for its performance in this regard, and outlining changes that the company would make [6]. Introduction of novelty to TikTok’s For You feed is a natural way to combat this negative press coverage. Since an instrumental value of novelty to users is exposure to other cultures, races, and locations, an instrumental value of novelty to TikTok would be introduction of users to novel lifestyles and ideas as a way to reduce the negative press that they are currently receiving.

The relevant harm of commission of implementing a novelty bias is the possibility of receiving bad PR if users perceive it as a deflection of blame. Users could potentially misinterpret the intentions of a novelty bias in TikTok, so we must allow for the possibility that it would be regarded as a stunt for positive PR. The harm of omission of failing to implement a novelty bias would be continued bad press with regards to race, and corresponding loss of active users and creators due to backlash. Without changes, the bad press coverage that TikTok is currently receiving will continue. There are reports of users being motivated to leave, and this backlash could reach a point that results in significant decreased engagement and fewer app downloads. As previously described, such changes would be highly detrimental to TikTok’s core interests. Bad press can also permanently tarnish a company’s reputation if it becomes recurring due to a company’s failure to institute change. Using novelty as a means to instrumentally combat bad press and solve racial concerns associated with the app is within both the power and interests of TikTok as a company.

4 All-things-considered Recommendation

In order to come to a conclusion with regards to TikTok’s potential moral obligation to employ an algorithmic novelty bias, we must weigh the different stakeholder values against one another, based on the moral importance of the underlying interests. In the simplest world, an algorithmic novelty bias would either be a net positive for all stakeholders, or a net negative for all. In either case, the ethical recommendation would be clear. However, it is very unlikely that a novelty bias will be uniformly positive or negative for all parties, so we must determine which interests are most morally important.

A few moral weights can be generally agreed upon. We contend that, since people have the ability to opt out of using TikTok as a social media platform, TikTok arguably does not have an obligation to harm themselves as a company with a novelty bias just to appease users. That is, if the proposed novelty bias violates the interests of TikTok, then it need not be implemented. A guiding principle in this regard is that users and creators are not being forced or coerced into using TikTok, and so the company is free to support its interests. Some social media platforms are arguably so large and ubiquitous that people cannot

opt out from using them (e.g., Facebook or Google, particularly if we include their analytics and authentication modules, as those run on many third party websites). However, TikTok does not occupy this kind of privileged role or status; neither users nor creators are being coerced (even implicitly) into using TikTok. Hence, TikTok does not bear any extra ethical obligations towards users beyond those implied by the company's own core values and interests.¹

However, the situation is more complex than suggested in the previous paragraph, partly because TikTok's own interests are dependent on user satisfaction in the long run. That is, we cannot understand TikTok's values without also understanding its users' and creators' values. In order to keep social engagement high and continue receiving downloads, TikTok's primary interest is to provide users and creators with (perceived) value. If users and creators are concerned about the app, whether from bad PR, lack of engaging content, inability to receive likes and views, or some other reason, then they will presumably opt out of using the service, thereby costing TikTok money and resources. Since users have this choice, and since TikTok molds their algorithm around user satisfaction, the company's interests blur into those of its users and creators.

In particular, although TikTok's values should dominate those of its users and creators in determining its moral obligations with regards to a novelty bias, we find that the values of those stakeholders are not opposed to one another. TikTok's greatest interests with regards to novelty bias are to increase user engagement and avoid bad press, where the latter is important because bad press tends to decrease user engagement by both current and potential users. These two user-grounded interests can pull in opposite directions.

Suppose that a novelty bias causes a decrease in user engagement, as would occur if the open empirical question about user values were answered in the negative. In this case, TikTok faces a tension between its interests: lower engagement through novelty bias, or increased risk of negative PR (so lower engagement) if they do not use a novelty bias. We conjecture, though unfortunately no empirical data are available, that the direct lower engagement from a novelty bias would outweigh user attrition due to bad PR.² Hence, TikTok should increase (direct) user value by not including a novelty bias.

In the other direction, suppose a novelty bias leads to an increase in user engagement (i.e., the user-centric open empirical question is answered in the affirmative). In that case, both of these interests can be satisfied by use of a novelty bias, though they can

¹ One might object to our use of a "no extra moral obligations to users/creators because no coercion of users/creators" principle here. We recognize that there are potential concerns about such a principle; for example, coercion is not a simple binary relation (present/absent). Even if there is no strong coercion, there may nonetheless be additional weak, partial, or defeasible moral obligations. However, our argument below will actually conclude that TikTok almost surely has the same obligations as if we rejected this principle, albeit for different reasons than if we believed that they have extra moral obligations to their users/creators. Hence, we do not further consider objections to our principle.

² This conjecture derives from our conjecture that people's decisions about TikTok use are based more in their personal enjoyment (or not), rather than the PR that TikTok receives in media, blogs, and so forth.

lead to distinct harms. In particular, if people perceive the implementation of a novelty bias as a superficial mechanism to avoid blame for the social justice concerns previously associated with TikTok's For You algorithm, then the company could receive even more backlash than before. Since the harms of implementing a novelty bias for the sake of reducing bad press outweigh those that result from increasing user engagement, the latter value should be valued more heavily.

Hence, we come to the conclusion that (unless bad PR reduces user engagement *much* more than novelty bias increases it) TikTok ethically ought to include an algorithmic novelty bias if and only if users and creators gain value from a novelty bias. That is, the blurring of the line between TikTok's values and those of its users and creators is nearly total. TikTok ethically ought to do whatever its users and creators value, where that last part is primarily dependent on the impact of a novelty bias on users' values. Since we think that it is likely that these groups value at least *some* novelty, we conclude that TikTok should aim to increase novelty and diversity in the For You feeds.

In contrast with many arguments that social networks and content platforms have an ethical obligation to increase diversity, our argument does not assume that TikTok has any special moral obligations to broader society, but only moral obligations to act in its own best (corporate) interests. And although TikTok did not initially appear to be ethically obligated to prioritize the interests of its users and creators (since those individuals could opt out), the company actually does bear an ethical obligation to support them, since that actually serves TikTok's own interests. Since user engagement is valued so heavily by TikTok, it is critical that the interests of users are satisfied in TikTok's mission. The precise (ethical) policy recommendations for TikTok must await answers to the two open empirical questions about whether a novelty bias provides value to users and/or creators, but we see that companies and their users need not stand in conflict with one another.

REFERENCES

- [1] Katie Elson Anderson, 2020. Getting acquainted with social networks and apps: It is time to talk about TikTok. *Library Hi Tech News*, 37, 4, 7-12.
- [2] Anja Bechmann and Kristoffer L. Nielbo, 2018. Are we exposed to the same 'news' in the News Feed?: An empirical analysis of filter bubbles as information similarity for Danish Facebook users. *Digital Journalism*, 6, 8, 990-1002.
- [3] Sam Biddle, Paulo Victor Ribeiro, and Tatiana Dias, 2020. Invisible censorship: TikTok told moderators to suppress posts by "ugly" people and the poor to attract new users. *The Intercept* (March 16, 2020).
- [4] Komal Kapoor, Vikas Kumar, Loren Terveen, Joseph A. Konstan, and Paul Schrater, 2015. "I like to explore sometimes": Adapting to dynamic user novelty preferences. In *Proceedings of the 9th ACM Conference on Recommender Systems (RecSys '15)*. ACM Press, New York, NY, 19-26.
- [5] Megan McCluskey, 2020. These TikTok creators say they're still being suppressed for posting Black Lives Matter content. *Time* (July 22, 2020).
- [6] Vanessa Pappas, 2020. A message to our Black community. <https://newsroom.tiktok.com/en-us/a-message-to-our-black-community> (June 1, 2020).
- [7] Siddharth Venkataramakrishnan, 2020. TikTok becomes political platform ahead of US election. *Financial Times* (June 1, 2020).