We would like to thank the reviewers for carefully reading our paper, and for their insightful and constructive comments. We will address these comments and update our draft accordingly. Please find our answers below.

Reviewer 1:
1) Comparison to other work and the no-repetition constraint. The main novelty of our paper lies in the fact that we consider various settings with the no-repetition constraint and yet we are able to derive tight regret lower bounds and to devise algorithms achieving these limits.
Note that the two papers, Zhou and Brunskill, IJCAI 2016, and Kawale et al, NeurIPS 2015) you suggested, *do not* consider the no-repetition constraint. To the best of our knowledge, Bresler et al. paper [3] and [4] are the only papers who theoretically analyze the no-repetition constraint (but with different models and an analysis that is not as exhaustive as ours, i.e., providing regret lower bounds and achieving algorithms).
2) Statistically identical users. In the first two models, we consider statistically identical users. Indeed this simplifies the analysis, but the latter is still challenging when accounting for the no-repetition constraint. Statistically identical users can be a good assumption when the budget T is rather small (if a user is not seen many times within this time budget). In addition, the first model also serves as a 'warm-up' to understand the last model, where we consider clustered users. Comparing the results of Models A (identical users) and C (clustered users) is insightful: it can tell us in which regimes exploiting the user clusters can be beneficial. For example, we remark that for small budgets (e.g. $T = o(\frac{m}{\Delta})$), we cannot leverage the user clusters to reduce regret.
3) Appendices. The proofs and pseudo-codes are in appendix due to space constraints. However, the main ideas of the algorithms are thoroughly discussed in the paper. Also note that when published on the NeurIPS website, both the main paper and the supplementary material are available.

Reviewer 2: For generic remarks on the novelty of our framework, please refer to answer 1) to Reviewer 1.
1) Novelty of the analysis. In Theorem 1, the lower bounds $R_{\mathrm{nr}}(T)$ and $R_{\mathrm{sp}}(T)$ can be easily derived using existing techniques in the bandit literature. The lower bound $R_{\mathrm{ic}}(T)$, however, is extremely challenging to derive due to the no-repetition constraint, which makes the problem non-asymptotic by nature (new items must be tested continuously). This term captures the fact that the decision-maker does not know the item clusters. We are unaware of any paper deriving regret lower bound due to such a lack of knowledge. Our main proof ingredient is a mapping of this lack of knowledge to a 2-arm bandit problem (see Lemma 7 Appendix D); and we find this argument elegant. We then use finite-time regret lower bound for this type of bandits. Importantly, $R_{\mathrm{ic}}(T)$ *does not* arise because of our definition of regret. In Appendix C, we prove that the difference between our definition of regret and the true regret is negligible compared to $R_{\mathrm{ic}}(T)$. Hence this term is not artificial at all.
2) About Model B. Our results are not simple extensions of those for the satisficing bandit problem, again due to the no-repetition constraint. Also, remember that users arrive randomly. Deriving the regret lower bound in Theorem 2 is again technical.
3) Dependence in $\frac{1}{\Delta}$. This dependence is present in our results (see the summary in the introduction). For example, in Theorem 1 (lower bound), this term is present in $R_{\mathrm{ic}}(T)$ through ϕ and in $R_{\mathrm{sp}}(T)$ explicitly. In Theorem 4, it is present in the first term of the regret upper bound, because in the sum, the term for $k = 2$ scales as $1/\Delta$. In Theorem 5, the term ε corresponds to the gap, and in Theorem 6, various gaps are also present.
4) Related work. Thanks for pointing out the two papers on low-rank bandits (again, these papers do not account for the no-repetition constraint). Regarding Bresler et al. 2014 paper, the setting is a bit different, even from our Model C (they do not have clustered items). Another important difference is that they consider that users arrive simultaneously, whereas, in our setting, users arrive sequentially, which is more realistic. Bresler et al. uses a cosine similarity approach that requires an ε-greedy exploration approach. In turn, this would lead to higher regret. Overall, we found it very difficult to compare our results to theirs.

Reviewer 3:
1) Changes in environment. Deriving efficient ETC algorithms was very already challenging. Devising algorithms that could adapt to non-stationary settings is left for future work.
2) About Model B. We wished to include this model to account for the possibility of un-clustered items.
3) Optimal regrets are independent of n. As part of our results, we figured out that under optimal algorithms, the regret does not scale in n for most regimes of (T, n, m). Note that in our regret expressions, $\frac{T}{m}(\ll n)$ roughly corresponds to the number of items that are recommended.
4) Choice of the title. The gaps between the upper/lower regret bounds are small. But we can modify the title to remove 'minimal'.

Reviewer 4: For the relevance of our Models A and B, please refer to answer 2) to Reviewer 1.
Assumption $\log(m) = o(n)$. Note that m can be polynomial in n. This is the same assumption as that made and justified in Bresler et al., NeurIPS 2014. Refer to the latter paper for an explanation of why the assumption is without loss of generality.