Human-AI Collaboration with Misaligned Preferences

Jiaxin Song[1], Parnian Shahkar[2], Kate Donahue[1,3], and Bhaskar Ray Chaudhury[1]

[1]University of Illinois, Urbana-Champaign
[2]University of California, Irvine
[3]Massachusetts Institute of Technology
{jiaxins8,kpd,braycha}@illinois.edu, shahkarp@uci.edu

Abstract

In many real-life settings, algorithms play the role of assistants, while humans ultimately make the final decision. Often, algorithms specifically act as curators, narrowing down a wide range of options into a smaller subset that the human picks between: consider content recommendation or chatbot responses to questions with multiple valid answers. Crucially, humans may not know their own preferences perfectly either, but instead may only have access to a noisy sampling over preferences. Algorithms can assist humans by curating a smaller subset of items, but must also face the challenge of *misalignment*: humans may have different preferences from each other (and from the algorithm), and the algorithm may not know the exact preferences of the human they are facing at any point in time. In this paper, we model and theoretically study such a setting. Specifically, we show instances where humans benefit by collaborating with a misaligned algorithm. Surprisingly, we show that humans gain more utility from a misaligned algorithm (which makes different mistakes) than from an aligned algorithm. Next, we build on this result by studying what properties of algorithms maximize human welfare, when the goals could be either utilitarian welfare or ensuring all humans benefit. We conclude by discussing implications for designers of algorithmic tools and policymakers.

1 Introduction

In recent years, advances in artificial intelligence and machine learning have become increasingly integrated into our daily lives: algorithms help suggest movies for us to watch, routes for us to drive on, or even generate novel content for us to rely on. However, outputs of algorithmic tools in this context almost never have the "final say": for example, while an algorithm can suggest multiple movies, the human makes the final decision on which movie she ultimately wants to watch. As a result, we often care about studying the performance of the human-algorithm *system*, rather than the performance of the algorithm in isolation. Human-algorithm systems have been studied extensively in more applied contexts, and in recent years, a more formal theoretical study of human-algorithm systems has grown (see Section 3 for more discussion).

In this paper, we focus on a specific instance of human-algorithm collaboration: where the algorithm acts as a *curator*, and the humans are *noisy* and potentially *misaligned* with each other and with the algorithm. Informally, an algorithm acts as a *curator* when its role is to narrow down a larger set of items to a smaller set, among which the human picks her favorite. This type of role is one of the most common in human-algorithm systems: e.g., in content recommendation, search, and some types of categorical prediction[1]. A human is *noisy* if she has inherent randomness in her ability to recognize her true preferences over items. "*To err is human*", and it has been widely recognized that humans are often imperfect at picking the "correct" item (e.g., see [Agranov and Ortoleva, 2017, Plackett, 1975, Luce et al., 1959, Hey and Orme, 1994]). A human is *misaligned* with another agent (human or algorithm) if they have different ground-truth preferences over items, which is distinct from their potentially noisy *realizations* of those preferences. For example, if Alice prefers horror movies and Bob prefers comedies, then their ground truth preferences over a set of movies would be *misaligned* with each other. Additionally, given a single algorithmic recommender, either Alice or Bob (or both) would be *misaligned* with the algorithm. Of course, in the

[1]For example, when a user requests directions, Google maps returns a small set of routes, and the user typically picks her final route from those routes (Figure 1).

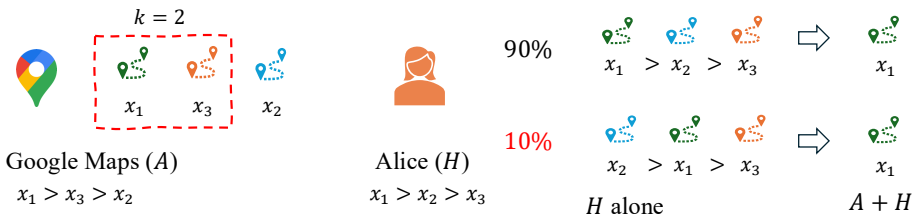


Figure 1: Human-AI collaboration between Google Maps (Algorithm) and Alice (Human). Here, we assume that the algorithm is deterministic, but the human is noisy and only the best item (x_1) has value. If Alice picked by herself, she would pick x_1 90% of the time, but when the algorithm deterministically reduces her set to $\{x_1, x_3\}$, she picks x_1 always.

limiting case of perfect personalization, both could have separate algorithms that are perfectly aligned with their preferences, but in realistic settings where one algorithm must serve a diverse range of humans, misalignment is a reality of life. Note that our framework has strong connections to other bodies of work, such as pluralistic alignment and conformal prediction: we discuss such connections in Section 3.

One of the key points of our paper is that there are settings where misalignment can be *helpful*. As a stylized example, consider Figure 1, which illustrates a collaboration between Google Maps (algorithm) and Alice (human). Google Maps recommends $k = 2$ routes $\{x_1, x_3\}$ to Alice from three routes $\{x_1, x_2, x_3\}$. The true preference of Alice is (x_1, x_2, x_3), but she cannot always identify her preference exactly (potentially because she only has access to imperfect information about the outcomes) and has a probability of 0.1 of wrongly ranking them as (x_2, x_1, x_3). As a result, she has a probability of 0.9 to choose her favorite route x_1 when making the decision alone. However, when following the suggestions of Google Maps, she is always able to rank x_1 before x_3 and choose the best route for her. The high-level goal of our paper is to study when settings like those in Figure 1 occur and when a human benefits from relying on a potentially misaligned algorithm.

Our contributions. In this work, we begin by formally defining our model in Section 2: this includes a description of how the human and algorithm interact, as well as the types of noise models that the human and/or algorithm exhibit, and gives a formal definition of misalignment. Additionally, we define several key objectives for the human-algorithm system to satisfy. For example, one objective is maximizing social welfare: if we assume each person has some utility for each item, expected (utilitarian) social welfare gives the total expected utility of all humans using the same algorithmic curation tool. Other objectives relate to the comparative performance of the human-algorithm system and the human by herself: does adding the algorithm increase the human's expected welfare, or decrease it? We say that a system achieves *uplift* if it increases the utility of every human, relative to what they would have achieved by themselves. Throughout the rest of this work, our goal will be to study when each of these objectives can be satisfied. We assume that an algorithm can be modified by either changing its ground-truth ordering over items (i.e., changing the training objective), or changing the noise level of the algorithm (i.e., changing the temperature in a stochastic model).

Our first problem focuses on when misalignment is helpful for an individual human:

Problem 1 (Comparative Benefits of Misalignment)**.** *Given a fixed noisy human, which algorithm leads to the greatest utility for the human?*

In Section 4, we begin by studying a fixed human who is considering a range of algorithmic tools that are all misaligned to varying degrees with the human. First, we note that it is possible for a human to benefit when using a misaligned algorithm: this should be intuitive to anyone who has successfully gained value out of an algorithmic tool that is imperfectly calibrated to their own preferences. Next, this section provides a comparative analysis of which type of algorithm assists the human more or less. Specifically, we show that increased alignment with the algorithm does *not* guarantee increased performance, and formally describe when misalignment may be helpful for utility. To give intuition, misalignment leads to "*different kinds of mistakes*" which can sometimes be helpful, if the mistakes are not too costly and if it makes it easier for the human to correctly identify their preferred items.

Next, Section 5 builds on this framework to answer the question, "How should we design an algorithm system to benefit a *population* of human users"?

Problem 2. *Can we find a ground-truth ranking and accuracy parameter such that the joint system maximizes utilitarian social welfare? Can we find settings such that we achieve uplift, or is such a setting impossible?*

Our main contributions can be summarized (informally) as follows,

- An algorithm that maximizes utilitarian welfare is always noiseless. While it is in general computationally hard to find this algorithm, we can design a mixed integer program (MIP) which is similar to the MIP used for assortment optimization under the Mallows model [Désir et al., 2016].
- However, we show cases where uplift can only be achieved by a noisy algorithm, aligning with the broader literature on the importance of randomization for fairness goals and providing motivation for the default non-zero temperature settings that many generative tools use.
- We provide efficient algorithms to validate whether a given algorithm achieves uplift, and identify natural sufficient conditions for when we can design algorithms that always achieve uplift.

While most of our contributions are theoretical, in each section we explore generalizations of our results through numerical simulations, showing that our main contributions generalize substantially beyond our formal model. Finally, in Section 7 and Section 8, we conclude by summarizing our main contributions, discussing extensions, and highlighting impacts for policymakers, platform designers, and applied researchers. At a high level, we view our work as highlighting an important tension between the desiderata of creating algorithms that *think like us* and *provide additional value* over what humans could do by themselves.

2 Model

2.1 Human-algorithm Collboration Model

In a human-algorithm collaboration model, there are two actors: an algorithm (A) and a human (H). There is a set of items $M = \{x_1, \ldots, x_m\}$ representing different outcomes (e.g., labels in prediction tasks, generative model output). The setting we study is that of where the algorithm acts as a *curator*, picking a subset of size k from a larger set of items of size m. The specific mode of interaction we assume is where the algorithm picks their top k from some noisy ranking over items, and the human picks their favorite among that set, according to their own noisy ranking over items[2] This setting may occur in settings such as content recommendation (e.g., Netflix or Spotify) or algorithmic assistants in predictive tasks.

These items can have different values to the actors. One notable special setting is called **top item recovery**, where only a single item has positive value. For example, these settings where the goal of the human is to recover the single best item, such as a job candidate who is most suitable, or the "best" route between two locations.

We say the algorithm and the human are *aligned* if they agree on the values of each item. However, as two actors may get information from different sources, perfect alignment cannot always be achieved. We consider *misalignment* within the actors through the following two aspects: either between the algorithm and the human, or between different humans. Formally, the humans come from a population of n types, where the ith type has probability p_i to occur. If human and algorithm are misaligned, but their top items are the same, we say they are *top-aligned*; otherwise, they are *top-misaligned*. We assume that human has descending values according to their ground-truth rankings. Let $v_{i,j}$ denote the value of the j-th best item for a human of type i. For each human of type $i \in [n]$, let her utility function be u_i, where $u_i(x) = v_{i,j}$ if item x is the j-th item in her ground-truth ranking ρ_i^*. We denote by x_H^i the item chosen by the human of type i when acting alone, by x_A the item chosen by the algorithm when acting alone, and by x_C^i the item chosen by the human of type i under collaboration.

2.2 Noisy Ranking Distribution

We assume that each agent can only access a noisy permutation over items: this may be because the agent only has access to imperfect or noisy data about each alternative, for example, or due to features such as temperature in generative AI tools [Ackley et al., 1985]. Denote the sampled rankings of the algorithm and human as $\pi \sim \mathcal{D}_a$

[2]Note that this is the same modeling assumption as [Donahue et al., 2024] in the unanchored case, but our results significantly generalize over their results by allowing agents to be misaligned.

and $\rho_i \sim \mathcal{D}_h^i$ respectively, where $\mathcal{D}_a$ concentrates at ground-truth ranking π^* and $\mathcal{D}_h^i$ concentrates at ρ_i^*. Denote by $x_i \succ_\pi x_j$ that x_i precedes x_j in π. Denote by $\pi(i)$ and $\pi[:k]$ the i-th item and the first k items (viewed as an unordered set) in a ranking π. Denote by $\pi \circ (x_i x_j)$ the ranking after swapping the locations of x_i and x_j in ranking π. All distributions in the paper are *inversion-monotonic* and *label-invariant*. Informally, adding inversion to a ranking only decreases its probability, and the distribution is not tied to the labels of the items. Both Mallows model [Mallows, 1957] and Plackett-Luce model [Luce et al., 1959, Plackett, 1975] satisfy the two properties.

Definition 1 (Inversion-monotonicity and label-invariance)**.** A distribution $\sigma \sim \mathcal{D}(\sigma^*)$ with ground-truth ranking σ^* is *inversion-monotonic* if $\mathbb{P}[\sigma = \sigma_1] > \mathbb{P}[\sigma = \sigma \circ (x_i, x_j)]$ where $x_i \succ_{\sigma^*} x_j$ and $x_i \succ_{\sigma_1} x_j$ and *label-invariant* if relabeling the items (i.e., applying an arbitrary permutation to all items) does not change the probability of any ranking. We say that two distributions are *isomorphic* if one is identical to the other by relabeling the items.

Mallows Model. The Mallows model [Mallows, 1957] generates a distribution over permutations based on their number of inversions relative to a central ranking (also known as the Kendall-Tau distance). A Mallows model $\mathcal{D}(\pi^*, \phi)$ consists of a central ranking π^* and an accuracy parameter $\phi \geq 0$. The probability of a permutation π occurring is $\mathbb{P}[\pi] = \frac{1}{Z} \cdot \exp(-\phi \cdot d(\pi^*, \pi))$, where Z is the normalization constant and $d(\pi^*, \pi)$ is the Kendall-Tau distance between π and π^*. We assume all accuracy parameters used in this paper are *positive*.

Plackett-Luce Model. The random utility model (RUM) is commonly used as a model of permutations [Thurstone, 1927], where a permutation is generated as follows: i.i.d. noise is added to the true value of each item to produce $\hat{v_i} = v_i + \epsilon$, for $\epsilon \sim \mathcal{D}$: then, the items are sorted in order of the noised values $\{\hat{v_i}\}$. Note that in a RUM, the probability of sampling a particular ranking depends not only on the number of inversions but also on the specific valuations of the items. Although there is no definition of "central ranking" in RUM, we refer to the ranking consistent with the item values as the central ranking. When ϵ follows i.i.d. Gumbel noise $G(\mu, \beta)$, following the analysis in [Luce et al., 1959, Plackett, 1975], the probability of a permutation σ is given by

$$\mathbb{P}[\pi] = \prod_{j=1}^{m} \frac{\exp(v_{\pi_j}/\beta)}{\sum_{\ell=j}^{m} \exp(v_{\pi_\ell}/\beta)}, \tag{1}$$

where v_{π_i} is the value of the i-th item in the ranking π. In particular, as β increases and item values become more heterogeneous, the probability of sampling a ground-truth ranking also increases.

2.3 Welfare objectives

We are interested in how an algorithm performs over a *population* of people who are all using the algorithm simultaneously, that is, *welfare objectives* over different humans who may receive different utility from the same algorithm. There are two main welfare objectives we will consider: utilitarian social welfare and uplift. Utilitarian social welfare is equal to the sum of the human's expected utilities (i.e., $\mathbb{E}[u_i(x_C^i)]$) weighted by the probabilities of every type of human (i.e., p_i). Formally,

Definition 2 (Expected utilitarian social welfare)**.** The *utilitarian social welfare* of a joint system is the sum of expected utility of the n types of humans $\sum_{i=1}^{n} p_i \cdot \mathbb{E}[u_i(x_C^i)]$.

Further, we are interested in whether the human-AI collaboration can lead to improvements in human utility, compared to the utility the human would experience if she solved the task herself. If an algorithm can simultaneously benefit a population of people in this way, we say that the algorithm achieves *uplift*.

Definition 3 (Uplift)**.** A human-algorithm joint system achieves *uplift* if the expected utility of every type of human is improved: $\mathbb{E}[u_i(x_C^i)] > \mathbb{E}[u_i(x_H^i)]$ for any $i \in [n]$[3].

3 Related Work

Human-algorithm collaboration. Our work relates to the general area of human-algorithm collaboration. In particular, there is a rich history of applied and empirical work in human-algorithm interaction: we refer interested

[3]We refer the reader to the related work section for discussion about the relation between uplift and complementarity [Bansal et al., 2021b]

readers to see [Preece et al., 1994, Kim, 2015, MacKenzie, 2024, Lazar et al., 2017] for textbook treatments. Specifically, our work relates to a growing literature using theoretical models to analyze humans interacting with algorithms. Some works study how to design algorithms to optimally assist humans [Bansal et al., 2021a, Chan et al., 2019, Donahue et al., 2022, Madras et al., 2018, Brown and Agarwal, 2022], including work incorporating models of human cognitive biases, such as [Li et al., 2024, Ibrahim et al., 2025, Chen et al., 2025]. Other work decides when human-algorithm teams perform well [Greenwood et al., 2024, Peng et al., 2024, Steyvers et al., 2022, Brown and Agarwal, 2024, Cowgill and Stevenson, 2020, Green and Chen, 2019, Alur et al., 2023, 2024, Guo et al., 2025], often relating to benchmarks such as complementarity (strict improvement over the human or algorithm alone defined in [Bansal et al., 2021b]). Some relevant literature reviews, taxonomies, and systematic studies in this space include [Gomez et al., 2025, Vaccaro et al., 2024, Rastogi et al., 2022].

Within human-algorithm collaboration, our work is most closely related to that of Donahue et al. [2024], which studies a similar setting where an algorithmic tool presents a top k subset to a noisy human. A key difference in our work is that we allow humans to be *misaligned* with the algorithm and study when misalignment is helpful. In fact, many of our contributions strictly generalize theirs, such as generalizing the utility function beyond top item recovery. Other related areas include *conformal prediction*, which studies how to optimize a subset of items (e.g., ensuring that the best item is presented with high probability): see [Angelopoulos et al., 2023, Fontana et al., 2023] for a summary of work in this area. Within this space, some works focus on optimizing the set of items that are presented [Straitouri et al., 2022, Wang et al., 2022, Straitouri and Rodriguez, 2023, De Toni et al., 2024, Babbar et al., 2022, De Toni et al., 2024, Hullman et al., 2025, Zhang et al., 2024], while others include more empirical analyses of specific settings [Angelopoulos et al., 2020, Arnaiz-Rodriguez et al., 2025]. In general, these works do not consider settings with multiple humans who may be misaligned with each other: one exception is [Corvelo Benz and Gomez Rodriguez, 2025], which studies an empirical setting on when algorithmic alignment is a helpful property for tools assisting human decision-makers. There is also a line of work studying human-AI collaboration through information elicitation [Steyvers et al., 2022, Corvelo Benz and Rodriguez, 2023, Corvelo Benz and Gomez Rodriguez, 2025, Collina et al., 2024, 2025a]. Steyvers et al. [2022] developed a Bayesian framework to combine the predictions of humans and AI algorithms with confidence scores. Corvelo Benz and Rodriguez [2023], Corvelo Benz and Gomez Rodriguez [2025] then showed that if the confidence value aligns with the human's confidence, the collaboration will benefit the human's decision-making. Collina et al. [2024, 2025a] studied the setting where each party holds different feature information and transmits their numerical prediction to each other at each round. Collina et al. [2024] generalizes Aumann's Agreement Theorem by relaxing the rationality assumptions and introducing calibration-based conditions on each party that ensure efficient convergence of the conversation. Building on this, Collina et al. [2025a] then proposes more communication-efficient protocols and removes the assumption of a common and correct prior. Finally, Collina et al. [2025b] models a Bayesian communication setting where a human is relying on information from multiple AI tools that are misaligned with each other (and with her) to make a discrete decision. Compared to their work, our work does not model the interaction of the human and the algorithm as a multi-round process. The algorithm acts more like a curator, instead of a rational agent.

Pluralistic alignment. Another research area that our paper relates to is pluralistic alignment. In this setting, the goal is generally to *align* an algorithm tool with users who have heterogenous preferences (e.g. [Sorensen et al., 2024]), especially the work connected with voting and social choice literature on aggregating diverse user preferences (e.g. [Conitzer et al., 2024, Dai and Fleisig, 2024, Shirali et al., 2025, A Alamdari et al., 2024, Ge et al., 2024, Pardeshi et al., 2024, Chen et al., 2024, Siththaranjan et al., 2023, Gölz et al., 2025]). One key difference between much of this work and our own is the meaning of noise. Often, work in this space (and voting in general) assumes that users know their own preferences and can represent them faithfully, which in the case of political opinions is often a fairly reasonable assumption. In our setting, we are assuming that humans may be able to only noisily access their true values over items and thus can benefit from the assistance of another agent (e.g., an algorithm). This is a more natural assumption in lower-stakes settings such as content recommendation or many types of content creation, such as writing code or daily emails. Separately, some papers study the performance of voting rules under noise (e.g., sporting competition), though they tend to be less focused on societal welfare objectives [Boehmer et al., 2022, Caragiannis et al., 2016].

Complementarity and Uplift. Note that the term of uplift is related to that of *complementarity* as defined in [Bansal et al., 2021b]: "*a human-AI system achieves complementary performance if it outperforms both the AI and the human acting alone*". As compared to complementarity, uplift is weaker because it only requires that the human-

algorithm system outperform the human alone, rather than the human and algorithm jointly. However, it also differs (and is weakly stronger) in that it is a *population-level* property: an algorithm must benefit all humans using it in order to achieve uplift. In addition, the application scenarios of complementarity and uplift also differ. The classical setting of complementarity usually assumes that both the human and the algorithm aim to predict an objective event. In contrast, our model does not assume the existence of an objective ranking of the items, and the goal of human-algorithm collaboration is to better align humans' preferences. Therefore, the performance of the algorithm alone is not well-defined in our setting and depends on how well the algorithm meets the human's preferences.

Other ranking/permutation models. The basic ranking model considered in our paper is defined over the number of inversions. A wide range of alternative permutation models has been studied in the literature. For example, the Bradley-Terry model [Bradley and Terry, 1952] and the Plackett-Luce model [Plackett, 1975, Luce et al., 1959] view a permutation as the consequence of a sequence of choices; the weighted Mallows model [Raman and Joachims, 2014] assigns weights to the inversions and defines the probability of a permutation by weighted Kendall-Tau distance. Moreover, Awasthi et al. [2014], Liu and Moitra [2018] studied learning a mixture of the Mallows model, which assumes the sampled permutation comes from a heterogeneous population.

4 Single Human with AI Assistant: Benefits to Misalignment

In this section, we begin by analyzing a single human who is considering multiple different algorithmic assistants. These algorithmic assistants differ from each other in their relative ground-truth orderings over items - e.g., some of them may be *misaligned* with the human's ordering over items. At first glance, we may expect the human's utility to decrease with increased misalignment with the algorithm. However, as mentioned by the previous example in Figure 1, misalignment sometimes better assists a human in making the decision. The goal of this section is to describe what those conditions are.

Without loss of generality, we relabel items by the human's ranking: $\rho^* = (x_1, \ldots, x_m)$. Meanwhile, we omit the index i for the human type and denote the human's value for the j-th item by v_j. All the proofs in this section can be found at Section B.

4.1 Top-item Recovery and Related Setting

As a warm-up, we first show a fine-grained analysis of the potential effect on the human's decision-making after swapping two specific items in the algorithm's ground-truth ranking.

Let x_i, x_j be two items with $i < j$, indicating that item x_i is (weakly) better than item x_j to the human. Denote algorithms A_1 and A_2 that are identical except that A_1 places x_i before x_j and A_2 inverts a pair of items x_i and x_j in their ground-truth rankings. We abuse the notations x_C^1 and x_C^2 for the items that the human picks when collaborating with A_1 (the aligned one) and A_2 (the misaligned one), respectively. We observed that,

Lemma 1. *For any item that is not i, the probability of picking that item is higher with the misaligned algorithm (and the probability of picking item i is lower). That is, for any $r \in [m]$ with $r \neq i$, $\mathbb{P}[x_C^1 = x_r] \leq \mathbb{P}[x_C^2 = x_r]$, and $\mathbb{P}[x_C^1 = x_i] \geq \mathbb{P}[x_C^2 = x_i]$*[4].

Proof sketch. We highlight the ideas here, and the full proof can be found at Section B. Since the misaligned algorithm A_2 places x_i lower than x_j in the ground-truth ranking compared to A_1, it is more likely to present x_r, x_j rather than x_r, x_i to the human. As x_j is relatively worse than x_i in the human's preference, the human is thus more likely to choose x_r when presented with x_r, x_j.

The itemwise probability comparison then implies Theorem 1 below, which distinguishes between misalignment on items that are valueless to the human and misalignment with items that are maximally valuable to the human, showing that the former are always helpful and the latter are always harmful.

Theorem 1. *The human gains **higher** utility with the misaligned algorithm when x_i and x_j are least valued, but **lower** utility when x_i is top valued.*

[4]Note that the two inequalities become tight only when the two items are completely indistinguishable to the human, i.e., swapping them does not change the probability of any ranking under the human's distribution. The following content considers the case where this condition does not hold.

We illustrate Theorem 1 in Example 1 below.

Example 1. *Consider the setting with three items, where the human's ground-truth ranking is $\rho^* = (x_1, x_2, x_3)$, and there are four potential algorithms (Figure 2). Suppose items x_2 and x_3 have equal value to the human. By repeated applications of Theorem 1, we can derive relationships between the value of each of these algorithms to the human: for example, Algorithm 2 is better than Algorithm 1 because it is created by an inversion in value-less items (x_2, x_3), and Algorithm 3 is better than Algorithm 4 by identical reasoning. Similarly, Algorithm 1 is better than Algorithm 4 because Algorithm 4 involves an inversion with the most valuable item (and Algorithm 2 is better than Algorithm 3 by identical reasoning).*

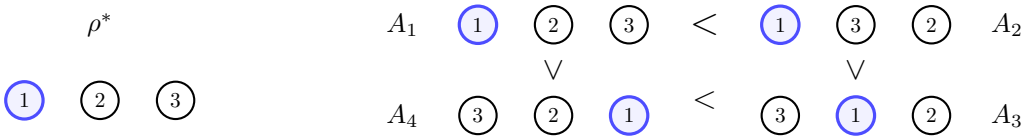


Figure 2: Illustration of Theorem 1 with 3 items, where the rounded node with number i represents item x_i and the most valuable item is in blue.

As a corollary of Theorem 1, Theorem 2 characterizes the best and the worst algorithm assistants for the top item recovery setting (where each human only has value for recovering her favorite item).

Theorem 2 (Best/worst strategy for top item recovery)**.** *In the top item recovery setting, the algorithm's ground-truth ranking π^* that maximizes human's expected utility is $\pi^* = (x_1, x_m, \ldots, x_2)$ while the one minimizing human's expected utility is $\pi^* = (x_2, \ldots, x_m, x_1)$.*

These results imply that Theorem 1 gives a *partial order* over which type of algorithms a given human would prefer. One natural question would be whether a *total order* over algorithms is possible. The following example illustrates why this is not always possible.

Example 2. *Consider the human working with two algorithm assistants A_1 and A_2, of which the ground-truth rankings are respectively $(x_1, \ldots, x_{10})$ and $(x_1, \ldots, x_5, x_{10}, \ldots, x_6)$. Note that the design of A_2 follows the insight of theorem 1 - aligning with the human on the high-valued items, while reversing the low-valued items. Figure 3 plots the relative performance between the two algorithmic assistants (A_1 and A_2) and the human working alone (H), where all of them follow the Mallows distribution, and the x-axis/y-axis are respectively the accuracy of the human and the algorithm. We assume that the two algorithmic assistants have the same accuracy. It can be noticed that the relative performance*

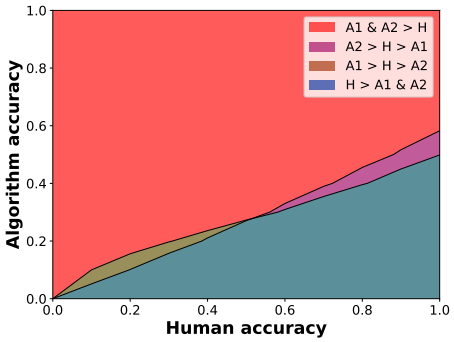


Figure 3: Comparison between human working with A_1, A_2, or alone, where A_1 & $A_2 > H$ denotes that the human performs better when assisted by either algorithm.

between A_1 and A_2 is not fixed and depends on the accuracy level.

	x_i (top), x_j (< top)	x_i, x_j (nearly top)	$j - i > \Delta$, x_i, x_j (< top)	x_i, x_j (bottom)
Highly accurate human	$A_1 > A_2$ (theorem 1)	$A_1 > A_2$ (lemma 2)	$A_2 > A_1$ (lemma 3)	$A_2 > A_1$ (theorem 1)
Accurate human	$A_1 > A_2$ (theorem 1)	$A_1 > A_2$ (lemma 2)	$A_2 > A_1$ (lemma 3)	$A_2 > A_1$ (theorem 1)
Highly noisy human	$A_1 > A_2$ (theorem 1)	$A_1 > A_2$ (lemma 2)	$A_1 > A_2$ (lemma 2)	$A_2 > A_1$ (theorem 1)

Table 1: Summary of the insights by Theorem 1 and its extensions. Note that A_1 is aligned and A_2 is misaligned, creased by swapping items x_i, x_j. x_i (top or bottom) indicates that x_i is the most (or least) valuable item to the human. $j - i > \Delta$ indicates that there is an index gap between item x_i and x_j. This table summarizes when misalignment is helpful or harmful.

4.2 Extensions to Specific Distributions

Another question is how Theorem 1 generalizes to settings where the inverted items x_i and x_j may not necessarily be the most or least valuable items to the human. This section investigates this question under two specific models: the Plackett-Luce model and the Mallows model by quantifying the probability changes in Lemma 1.

We extend Theorem 1 both theoretically and numerically: **(i)** We establish more general conditions under which misalignment is guaranteed to be either beneficial or harmful under the Mallows and Plackett–Luce models. These theoretical extensions provide a characterization of when swapping x_i and x_j is beneficial or not. We summarize the insights in Table 1. **(ii)** We present a numerical study in Section 9.1, examining scenarios where both the human and the algorithm follow either a Mallows model or a Random Utility Model (RUM). Our results indicate that the conclusions of Theorem 1 continue to hold even when x_i and x_j take relatively small values.

Let π_1 and π_2 respectively be the rankings output by A_1 and A_2. The following lemma first provides a sufficient condition when misalignment is harmful to the human.

Lemma 2. *Sufficient condition for when collaboration with a misaligned algorithm is* ***harmful*** *for the human under the Mallows model with accuracy parameter ϕ_h is $v_1 - v_i \leq \frac{\exp(-\phi_h \Delta)}{1-\exp(-\phi_h \Delta)} \Delta$ where $\Delta = j - i$ and $k = 2$.*

We now provide further intuition on when the above conditions hold. The condition is more likely to hold when the value of the most valuable item, v_1, does not significantly exceed the value of x_i, v_i, or ϕ_h is sufficiently small. The former condition means that x_i is nearly the top item, while the latter one indicates the human is *highly noisy*. Under these conditions, collaboration with the misaligned algorithm tends to only hurt the human.

It is worth noting that Lemma 2 does not provide a necessary condition: when Δ is large, the right-hand side tends to zero, and the condition is hard to satisfy. However, the human may still be harmed in this case. For instance, Theorem 1 shows that the human always receives less utility once x_i is the most valuable item while x_j is less valuable.

Takeaway 1. *The human gets hurt in working with the misaligned algorithm if either (i) the human is* highly inaccurate*, or (ii) both x_i and x_j are* nearly *top items.*

Next, we complement Lemma 2 by providing sufficient conditions when misalignment is beneficial to the human.

Lemma 3. *Sufficient condition for when collaboration with a misaligned algorithm is* ***helpful*** *for the human under the Mallows model is when some $i' \in [i-1]$ satisfies*

$$\frac{v_{i'}}{v_i} \geq \frac{\sum_{r \neq i,j} \psi(i,j,r)}{\sum_{r=1}^{i'} \psi(i,j,r) \exp(-\phi_h \cdot (j - r + 1))} \frac{1}{1 - \exp(-\phi_h \cdot \Delta)},$$

where $\Delta = j - i + 1$ and $\psi(i,j,r) = \mathbb{P}[\pi_1[:2] = \{x_i, x_r\}] - \mathbb{P}[\pi_1[:2] = \{x_j, x_r\}]$ is always nonnegative.

Notice that $\psi(i,j,r)$ on the right-hand side depends only on the algorithmic assistant's distribution. In the case where most of the values of $\psi(i,j,r)$ concentrate on terms with $r \leq i'$ (e.g., the algorithm is aligned with the human on the high-valued items), the first term asks for an exponential gap between $v_{i'}$ and v_i. Meanwhile, the second term depends on the index gap between the two swapped items (i.e., Δ) and the accuracy level of the human (i.e., ϕ_h). The condition is more likely to hold when the index gap and the accuracy level of the human are both non-trivial, i.e., larger than some non-negligible constants. The following provides a concrete example where the condition holds.

Example 3. *Consider the setting of four items and two algorithm assistants that randomly select two items from the top three of their ground-truth rankings, displaying the top two items with a higher probability than displaying other subsets. Suppose the ground-truth rankings of A_1 and A_2 are x_1, x_2, x_3, x_4 and x_1, x_3, x_2, x_4 respectively, where x_2 and x_3 are the misaligned items. The human has value of $v_1 = 100, v_2 = 2, v_3 = 1, v_4 = 1$. The human's ranking satisfies the Mallows model with accuracy parameter $\phi_h = 1$. Let $i' = 1$. Then the two sides of Lemma 3 are respectively given by*

$$LHS = \frac{v_1}{v_2} = \frac{100}{2} = 50 \quad RHS \leq \exp(3) \cdot \frac{1}{1 - \exp(-1)} < LHS\,.$$

Therefore, the human receives a higher utility when collaborating with the misaligned algorithm A_2 than the aligned one A_1.

Takeaway 2. *The human benefits from collaborating with a misaligned algorithm when there exists an item that is* significantly more valuable *than x_i and x_j, and (i) the human possesses a* non-trivial *level of accuracy, and (ii) there is a clear index gap between x_i and x_j in their ranking.*

Lastly, we remark that a similar set of extensions also holds under the Plackett–Luce model. When the human's distribution satisfies the Plackett-Luce model with noise parameter β, we have the following two results.

Lemma 4. *Sufficient condition for when collaboration with a misaligned algorithm is **harmful** for the human under the Plackett-Luce model is when $v_0 - v_j \leq 1.27\beta$.*

Lemma 5. *Sufficient condition for when collaboration with a misaligned algorithm is **helpful** under the Plackett-Luce model for the human is when some $i' \in [i-1]$ satisfies*

$$\frac{v_{i'}}{v_i} \geq \frac{\sum_{r \neq i,j} \frac{\psi(i,j,r)}{\exp((v_r - v_i)/\beta)+1}}{\sum_{r=1}^{i'} \frac{\psi(i,j,r)}{\exp((v_r - v_i)/\beta)+1}} \cdot \frac{2}{1 - \exp\left(-\Delta/\beta\right)},$$

where $\Delta = v_i - v_j$ and $\psi(i,j,r) = \mathbb{P}[\pi_1[:2] = \{x_i, x_r\}] - \mathbb{P}[\pi_1[:2] = \{x_j, x_r\}] \geq 0$.

At a high level, similar to Lemma 2, when the human is highly noisy or x_i and x_j are nearly top items, misalignment is harmful. Slightly different from Lemma 3, the positive result under the Plackett–Luce model considers the value gap between the two items (i.e., $v_i - v_j$), which is consistent with the definition of the distribution.

5 Multiple Humans Using Single AI: Welfare Objectives

In Section 4, we studied the question of which (potentially misaligned) algorithm an *individual* human would prefer. In this section, we study the question of which algorithm to select in order to satisfy welfare notions over a *population* of humans. In particular, this section focuses on the Mallows model (the rankings of both the algorithm and the human satisfy the Mallows model), and we will study algorithms that can either rearrange the orderings of items of its ground-truth ranking π^* or change its accuracy level ϕ_a.

There are various societal objectives that we may consider: in Section 5.1, we study expected utilitarian social welfare. For this objective, we show that finding the optimal arrangement of items (e.g., determining the optimal central ordering of the algorithm) is NP-hard. Finally, even though it is computationally hard, it can be formulated as a linear mixed-integer program (MIP) with only a linear number of binary variables. However, a welfare-maximized strategy may not always be desirable in every respect: next, Section 5.2 turns to other objectives, such as uplift. We provide counterexamples where maximizing welfare violates uplift, and we also find cases where introducing a higher noise level helps achieve uplift. Finally, we provide illustrations of solutions achieving uplift for special cases (such as top-item recovery). We conclude by illustrating generalizations of our results with numerical simulations.

5.1 Social Welfare Maximization

First, recall that our definition of expected utilitarian social welfare is given by:

Definition 2 (Expected utilitarian social welfare)**.** The *utilitarian social welfare* of a joint system is the sum of expected utility of the n types of humans $\sum_{i=1}^{n} p_i \cdot \mathbb{E}[u_i(x_C^i)]$.

Next, we explore how we can maximize the utilitarian social welfare. Our main result is Theorem 3, which shows that while selecting the optimal accuracy (or noise) parameter is extremely straightforward, selecting the optimal algorithmic ordering over items can be NP-hard. Finding the optimal arrangement of items can be reduced to the problem of finding the maximum independent set, which is known to be NP-complete. In addition, given an optimal arrangement, minimizing noise is always optimal. However, minimizing noise is not always better for any arrangement – especially if the arrangement is highly (equivalently, maximizing accuracy) misaligned with the ground-truth arrangements of the majority of the population.

Theorem 3 (Computational hardness)**.** *The expected social welfare is maximized when the algorithm's distribution is noiseless. However, it remains* NP*-hard to find π^* that maximizes the expected social welfare, even in the* top-item recovery *setting.*

To complement the hardness result, we show that the welfare-maximizing algorithm can be exactly characterized as the optimal solution to a mixed-integer program (MIP) and solved efficiently in practice, even for $m = 20$ items and heterogeneous populations of humans with 720 types of humans. The full proof and detailed empirical results can be found at Section C.1 and Section 9.2. Meanwhile, although finding the welfare-maximizing algorithm in the top-item recovery setting is computationally hard, one can still efficiently identify the welfare-maximizing algorithm among those that *achieve uplift*, as we will show in Lemma 8.

Finally, we briefly discuss the noiselessness result for the optimal arrangement in Theorem 3. This comes by showing that the objective can be rewritten as the sum over all possible sets of k items that could be presented to the population of humans: of all of those sets, one maximizes utilitarian welfare, and thus deterministically returning that set would maximize welfare as in Definition 2.

5.2 Uplift

However, a social welfare-maximizing strategy may not always be desired as it can end up hurting some users (i.e, they make worse decisions than they would acting by themselves). This is especially true if a certain type of human's ground-truth ranking ρ_i^* does not align well with most other humans' ground-truth rankings, and so to maximize expected social welfare, the algorithm may optimize for one type of human preference at the expense of other types of humans. As a result, these types of humans may receive less accurate recommendations from the algorithm in the collaboration: they may even obtain lower utility than they would by themselves. A social welfare desiderata relating to this objective is *uplift* across the population of people: when is it possible to design an algorithm such that every type of human benefits, relative to the utility they would obtain from solving the problem by themselves?

Definition 3 (Uplift)**.** A human-algorithm joint system achieves *uplift* if the expected utility of every type of human is improved: $\mathbb{E}[u_i(x_C^i)] > \mathbb{E}[u_i(x_H^i)]$ for any $i \in [n]$[5].

As a warm-up, we first consider the special case where the human is aligned with the algorithm on every item for which it has positive utility: the following lemma shows that uplift (e.g., strict benefit for the human) is guaranteed.

Lemma 6. *Suppose human of type i only has positive values for the top T items, i.e., $v_{i,j} > 0$ for $j \leq T$ and $v_{i,j} = 0$ for $j > T$. Then if $\pi^*(j) = \rho^*(j)$ for any $j \leq T$, $\phi_a = \phi_h$, and $1 < k < m$, uplift is achieved.*

Note that Lemma 6 strictly generalizes results in [Donahue et al., 2024], which provided these results only for when exactly 2 items are presented and the human is in the top item recovery case (only has value for her top item). Moreover, Lemma 6 does not put any constraint on the zero-valued items, and the algorithm and human can be misaligned on these items. The proof can be found at Section C.3.

5.2.1 Tension between Social Welfare and Uplift

In the above case, a social welfare maximizing arrangement implies uplift (if feasible) since there is only one type of human. However, when it comes to multiple misaligned humans, forcing welfare-maximization could hurt some type of human, e.g., when there are two types of humans with population shares of 0.99 and 0.01 and completely

[5]We refer the reader to the related work section for discussion about the relation between uplift and complementarity [Bansal et al., 2021b]

different preferences, maximizing welfare makes the algorithm fully aligned with the preferences of the majority type, thereby hurting the minority group.

Secondly, uplift can be viewed as a type of fairness notion: it requires that each type of human receive some benefit from using the algorithm. In contrast, social welfare is a well-accepted metric for evaluating the algorithmic efficiency. In Section 9.3, we further describe numerical experiments on connections between the two objectives. We study the change in optimal social welfare after enforcing the uplift constraint. Usually, enforcing uplift will reduce the optimal social welfare, especially when the distribution of humans is highly imbalanced, e.g., certain types of humans dominate the population. In addition, in Section 6, we compare the two optimal algorithmic solutions, each optimizing a single welfare objective, on a real-world dataset. We also find that solely maximizing utilitarian welfare could cause extreme unfairness.

5.2.2 Small Noise Facilitates Uplift

Given the tension between the two objectives, small noise may be an effective method to achieve uplift. Considering the same setting as the last example, it seems likely that a small degree of randomization could sharply increase welfare for the less populous type of human. Lemma 7 shows this more formally, and the proof can be found at Section C.4.

Lemma 7. *There exist settings where uplift can occur at lower accuracy (higher noise), but fails at higher accuracy.*

We remark that Lemma 7's message is in line with other works on fairness (e.g., [Jain et al., 2024, Singh and Joachims, 2018]) that deem randomization to be helpful for fairness. One difference in our setting is that we expect that in many cases there will be relatively large benefits in terms of uplift from relatively small levels of noise (and thus, relatively small impact on the utilitarian welfare objective). This is because small amounts of noise that result in occasionally including a new item has *asymmetric costs and benefits*: it has large benefits for users who stronger prefer that item to those that are already included because they will be able to select it with relatively high probability. However, it has small negative impacts on users who have low value for that item, since on the occasions when it is included, they will have a high probability of correctly ignoring that item as irrelevant.

5.2.3 General Results for Uplift

Motivated by the above examples, we consider how to design an algorithm that meets the objective of uplift. We begin by noting that there are cases where achieving uplift is impossible. To build intuition, consider a scenario in which humans are highly misaligned in their top preferences. In such situations, an uplift strategy is only feasible if all of their most preferred items can be accommodated within the limited window of size k. This naturally motivates characterizing instances where uplift is achievable. To this end, we first show that it remains NP-hard to determine the existence of a strategy (which may be noisy) that achieves uplift. However, we note that when the size of presented items k is constant, one can verify whether a given strategy achieves uplift. Thus, for a constant k, the problem is in NP. We note that in real-life settings, we expect k to be relatively small, given that humans would likely be unable to process very large subsets of items. The proof (can be found at Section C.5) is based on a reduction from the vertex cover problem.

Theorem 4. *It is* NP*-hard to determine whether there exists a strategy satisfying uplift (π^*, ϕ_a). Further, it is* NP*-complete when k is given as a constant (which means whether a given strategy (π^*, ϕ_a) achieves uplift can be verified in polynomial time).*

Next, we provide characterizations of what an uplift strategy looks like in some special cases, such as top item recovery (where we show that the optima algorithm is a noiseless algorithm and presents only items that at least one human values), and where the humans share the same set of best items.

For the top item recovery setting, consider a simple strategy that ranks all these top items first in the algorithm's ranking and sets the noise as zero (equivalently, set ϕ_a as $+\infty$). Lemma 8 shows that the simple strategy achieves uplift. We note that maximizing social welfare remains NP-hard even in this special case (using the same reduction of Theorem 3).

Lemma 8. *In the top item recovery setting, when the ground-truth rankings of humans satisfy $|\mathcal{M}_0| \leq m - 1$ where $\mathcal{M}_0$ is the set of distinct items that some human has positive value for, then always presenting $\mathcal{M}_0$ to the human achieves uplift and also maximizes the expected utilitarian social welfare among all the noiseless algorithms that achieve uplift.*

6 Empirical Study

In this section, we conduct an empirical study of human-AI collaboration with misaligned ground-truth rankings on a real-world dataset to answer the following question: *how does human accuracy affect the tension between different welfare objectives?*

We use the sushi preference dataset [Kamishima, 2003], which consists of about 5k rankings of ten kinds of sushi, where we denote the ten types of sushi by $x_1, \ldots, x_{10}$. For computational efficiency, we consider only the partial rankings of the humans in our sushi dataset over the items $x_1, \ldots, x_5$, treating these partial rankings as their ground-truth preferences. Each ranking is also associated with the fraction of the population in the dataset who had this preference over the types of sushi. A detailed breakdown of the most frequent sushi rankings is presented in Table 3 (Section D).

Note that in this dataset, it is inherently impossible to distinguish between a) population-level heterogeneity in preferences, and b) user-level noise in reporting those preferences. To our knowledge, there is not a preference dataset that cleanly displays both a) and b). As such, in this section we make the assumption that users have reported their sushi preferences perfectly (no errors from b) and that the preference dataset simply reflects user heterogeneity (only difference is from a). In order to model user-level noise, we add noise (e.g., Mallows) on top of the ground-truth preferences within [Kamishima, 2003]. Furthermore, because this dataset does not include explicit human valuation scores for the sushi items, we assume item values within each ranking is determined by their *Borda counts*. Formally, given a ranking π over m items $\{x_1, \ldots, x_m\}$, where $\pi(i)$ denotes the position of item x_i (with $\pi(i) = 1$ being the most preferred), the Borda score assigned to item x_i is

$$B(x_i; \pi) = m - \pi(i).$$

Thus, the top-ranked item receives a score of $m - 1$, the second-ranked item a score of $m - 2$, and so on, with the least-preferred item receiving a score of 0. This provides a simple cardinal approximation of ordinal preferences, which we use as the value function in our experiments.

The objective is to design an algorithm that presents only three items from $x_1, \ldots, x_5$ to each human to maximize a chosen notion of welfare. We find this experiment especially natural for our setting: if a restaurant created a "menu" of options that is quite large (> 100), a user would probably have a low probability of finding sushi she likes. However, if the restaurant creates a very small menu, then user-level heterogeneity in preferences might mean that some humans are hurt by the algorithm. In this section, we explore these trade-offs within the specific sushi preference dataset. We assume the algorithm is perfectly noiseless, so the menu always consists of the top three sushis in its ground-truth ranking. To study the tension between different welfare objectives, we compare three different welfare objectives:

1. Let A_w denote the algorithm that maximizes *utilitarian welfare*.
2. Let A_m denote the *majority ranking*, i.e., the ranking supported by the largest number of individuals.
3. Let A_u denote the algorithm that maximizes *uplift*, defined as the number of humans who benefit from collaboration.

In Figure 4, we plot the utilitarian welfare and uplift achieved by each of the algorithms across different levels of human accuracy. In the left plot, we observe that as humans become more accurate, their expected utilitarian welfare increases under both A_m and A_w. However, this monotonic improvement does not hold for A_u: beyond a certain human accuracy threshold, the performance of this algorithm declines.

Interestingly, in the right plot, we observe that as humans become more accurate, after a certain threshold, all algorithms can guarantee positive uplift for a smaller fraction of individuals. This suggests that when human judgments are noisier, a larger proportion of people benefit from collaboration with a fixed algorithm. However, as humans become more accurate, no single algorithm remains universally beneficial. Moreover, we observe that the algorithm optimizing utilitarian welfare performs substantially worse than the other two in terms of uplift across most accuracy regimes. This suggests a trade-off between the uplift and utilitarian welfare objectives. In this dataset, the majority algorithm performs relatively well on both metrics.

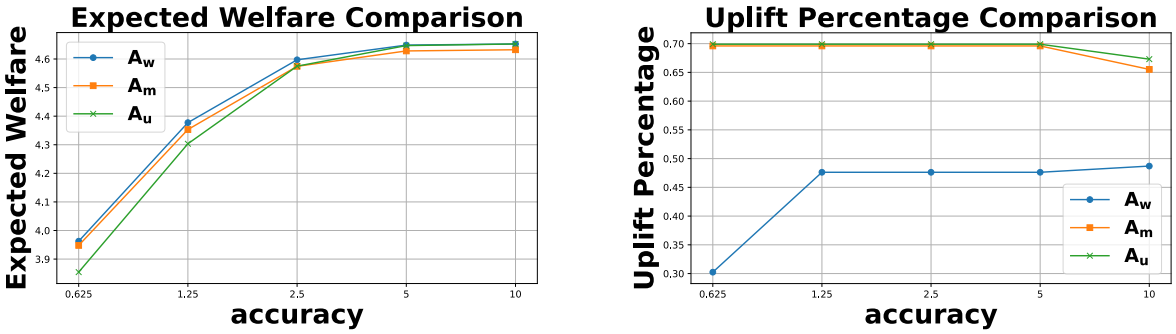


Figure 4: Utilitarian welfare and uplift achieved by the majority algorithm A_m, the welfare-maximizing algorithm A_w, and the uplift-maximizing algorithm A_u across varying levels of human accuracy.

7 Discussion

In this paper, we studied a model of human-algorithm collaboration in a setting where the algorithm acts as a *curator* of items for a *noisy* human, and the algorithm may be *misaligned* with the human's preferences over items. In particular, we define misalignment as a disagreement on the relative ordering over items by value: while such misalignment can be detrimental (because the algorithmic assistant returns items the human may find suboptimal), it can also be *beneficial* (because the assistant makes "different mistakes" that the human may be more easily able to correct). Our goal in this paper was to analyze when the relative benefits of misalignment outweigh the costs.

In Section 4 we began by studying the preferences an *individual* human would have over an algorithmic assistant, showing that certain types of misalignment are helpful. In Section 5, we turned to the question of designing an algorithm to maximize welfare over a *population* of humans who have diverse preferences. While in general it is not possible to simultaneously maximize the welfare of each human engaging with the algorithm, we study how to maximize some collective welfare objectives, like maximizing utilitarian welfare or ensuring that each human strictly benefits from using the algorithmic assistant, relative to what utility she could get by herself. In general, we show that exactly achieving either of these objectives is computationally hard. However, we also give positive results for restricted settings: For maximizing utilitarian welfare, we present a mixed-integer program, building on similar efficient formulations previously applied to problems involving Mallows preferences. In addition, we identify compelling special cases where uplift is attainable and provide an explicit algorithm that achieves it. Throughout, we supplement our theoretical results with numerical simulations, exploring generalizations in permutation distributions as well as values agents have for each item.

8 Extensions

There are numerous potential extensions to our work. In Section 8.1, we describe a few extensions of our theoretical model, including the extensions of the noisy model and the way of human-algorithm interaction. In Section 8.2, we discuss substantial implications for policy making, specifically on the relative benefits and harms of misalignment for users.

8.1 Model Extensions

First, some extensions could involve relaxing our theoretical model. For example, our main theoretical results focus on permutation models related to the number of inversions (e.g., the Mallows model). While we show in numerical simulations that our core results generalize to other models such as RUM, extending these results theoretically (or extending our results to other classes of permutations) could be interesting. Additionally, one core assumption of our work is that humans have the same *magnitude of preferences*: that is, every human has the same value for their favorite item, the same value for their second-favorite item, and so on (even though the identity of their favorite items may differ). This corresponds to assuming that all humans are equally "picky". While this assumption is necessary for tractability purposes, this is likely not true in general: some humans may be equally happy with

multiple options, while other humans may be much more selective. Relaxing this assumption may lead to different trade-offs in objectives.

Other extensions could change our model of human-algorithm interaction. For example, our model of subset curation assumes that the human observes the set of presented items as an unordered *set* and is not biased by any inherent ordering of items. This is almost certainly not true in practice, as it has been demonstrated that humans are often biased by the order in which items are presented (e.g., see [Joachims et al., 2007, Agichtein et al., 2006, Craswell et al., 2008]), and this effect has been studied in human-algorithm collaboration more specifically (e.g. [McLaughlin and Spiess, 2024, Donahue et al., 2024, Wang and Pananjady, 2022]). It could also be possible that the human and algorithm undergo a sequence of back-and-forth interactions which ultimately end up at a final solution (as in [Collina et al., 2025a]), which may lead to different implications on personalization. Finally, our results on optimizing the algorithm for different social welfare functions assume that the algorithm has perfect knowledge of the distribution of human preferences. In practice, this is almost certainly not the case, and it could be useful to study settings where the algorithm must learn human preferences over multiple interactions.

8.2 Implications for Policy

Our results have potential implications for policymakers, designers of algorithms, and users of algorithmic tools.

Potential for benefits in misalignment. In some real-life settings, humans have near-perfect ability to identify the "correct" outcome (e.g., the political opinion that they most closely identify with). However, in many real-life settings, humans may only be able to imperfectly (noisily) recover their own true preferences: prior research has shown that humans noisily pick the favorite item from sets, such as presented movies or items on a menu [Agranov and Ortoleva, 2017, Plackett, 1975, Luce et al., 1959, Hey and Orme, 1994]. In this setting, our results show that having algorithms that are "misaligned" in specific ways can be beneficial. This implies that designers of algorithms should maybe not always focus on designing algorithms that personalize to individual humans (eliminating misalignment): in settings where humans are noisy, personalizing may end up being harmful. This implies that algorithmic designers may want to take different strategies towards alignment depending on the context in which humans will use the tool, focusing on alignment more in noiseless settings and less in settings where humans are expected to know their own preferences less well.

Benefits of positive temperature. Our results also show that there can be benefits to algorithmic noise (e.g., positive temperature) when a single algorithm is serving a population of heterogeneous users. Benefits of randomization for goals related to fairness have been shown in other settings (e.g., [Jain et al., 2024, Singh and Joachims, 2018]), though to our knowledge, this question is less well studied in the context of complementarity or similar benchmarks. Most generative AI tools have a constant temperature across multiple queries: our work suggests that for certain welfare objectives, it could be helpful to have temperature change dynamically, being lower in settings where human opinions are roughly in agreement and larger in settings where human opinions are more diverse.

Implications for users. Our results also suggest strategies for users of algorithmic tools. Users sometimes have direct and indirect ways of controlling personalization: for example, they can choose to accept or delete cookies, can opt to use incognito mode, can volunteer or (sometimes) delete information generative tools have stored on their prior interactions, or can select tools that they believe to be more or less closely aligned with their beliefs. Our results provide motivation for users to sometimes strategically seek options that reduce personalization, even absent other concerns such as privacy.

Tensions in societal goals. It should not be surprising that some societal goals may be in tension with each other. However, we hope that our research could highlight particular objectives that are relatively understudied in the pluralistic alignment setting, such as ensuring all types of humans benefit in a randomized task (as compared to their utility if they solved the problem themselves). While we believe that there could be contexts in which either utilitarian welfare or goals like uplift would be optimal, we think it is worth having designers deliberately decide on which they may choose to optimize in different contexts.

9 Numerical Simulations

9.1 Numerical Extension of theorem 1

We consider $m = 4$ items, of which the algorithm displays $k = 2$. For the human, the value of item x_i, $v_i \propto e^{-\beta \cdot j}$, where $\beta \geq 0$ controls the heterogeneity of values. Heterogeneity helps to relax the top item recovery case smoothly.

Extensions to Mallows Model. The first experiment assumes that both human and algorithm rankings follow a Mallows model with the same accuracy of $\phi_a = \phi_h = 0.5$. Let $\rho^* = (x_1, x_2, x_3, x_4)$ be the human's ground-truth ranking, and let π^* be an arbitrary algorithm's ground-truth ranking. Figure 5 plots the *difference* in human's expected utility working with various misaligned algorithms, and the aligned algorithm one. The right figure Figure 5 restricted to top-aligned algorithms (algorithms that place x_1 first).

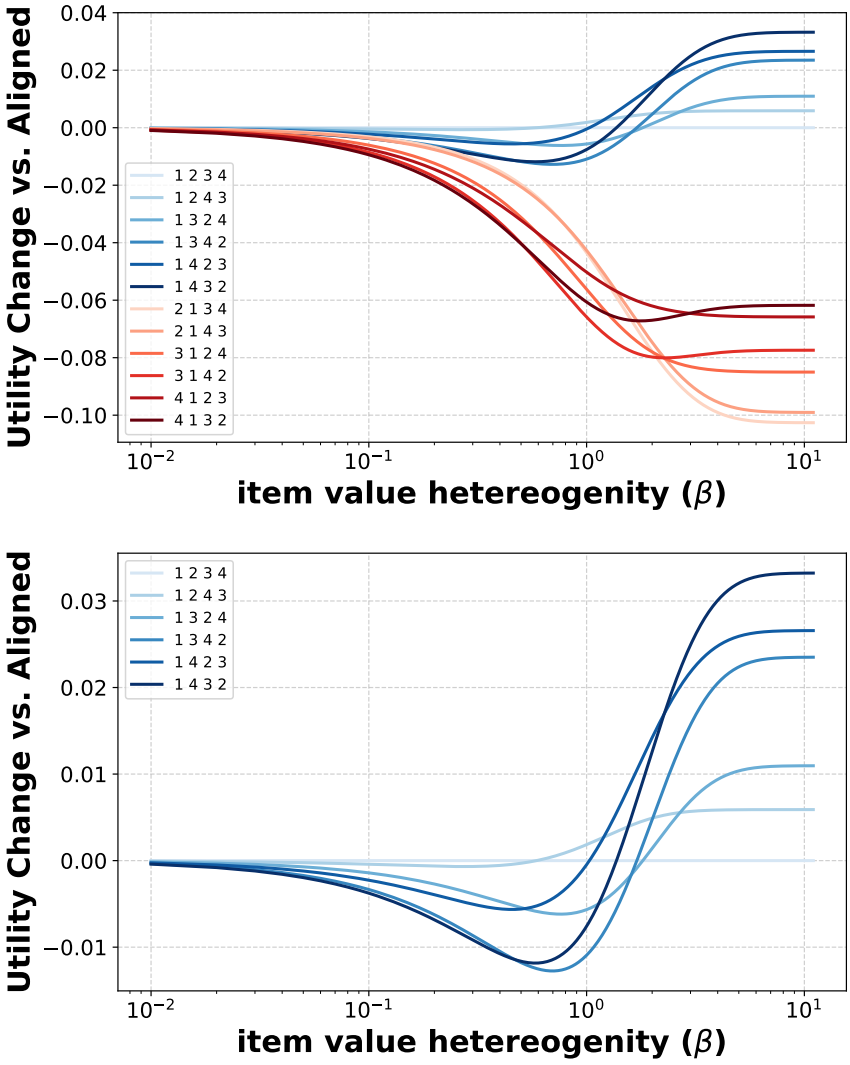


Figure 5: Comparison of human's expected utility differences after collaboration with a misaligned and an aligned algorithm, as a function of β. Each curve represents an algorithm with the corresponding ground-truth ranking.

For small β, valuations are nearly uniform, so the human's utility is similar across all algorithms. As β increases, the aligned algorithm performs best, but when β becomes sufficiently large, the utility is dominated by the top item x_1, and top-aligned algorithms yield higher utility. In the extreme top-item recovery case (Theorem 2), the

most inverted top-aligned algorithm achieves the highest utility, explaining why the curve for $\pi^* = (x_1, x_4, x_3, x_2)$ dominates at large β. Consistent with Theorem 1, all top-aligned algorithms outperform the aligned one when β is large.

Extensions to Random Utility Model. We perform a similar simulation by assuming that both the human and the algorithm follow the Plackett-Luce RUM with Gumbel noises of 0.1. In Figure 6, we observe that top-aligned algorithms quickly outperform others. This observation mirrors the second property established in Theorem 1: swapping item x_1 from the top position with any lower-ranked item yields a new algorithm ranking that is less advantageous for the human.

We emphasize an important distinction in the RUM setting: unlike the Mallows model, modifying item values alters the probabilities of sampling different rankings. As a result, when β becomes large, the probability that top-aligned algorithms present item x_1, as well as the probability that the human selects item x_1, both approach 1. This convergence explains why, at high β values, the expected utilities of all top-aligned algorithms and the aligned algorithm become nearly identical in RUM.

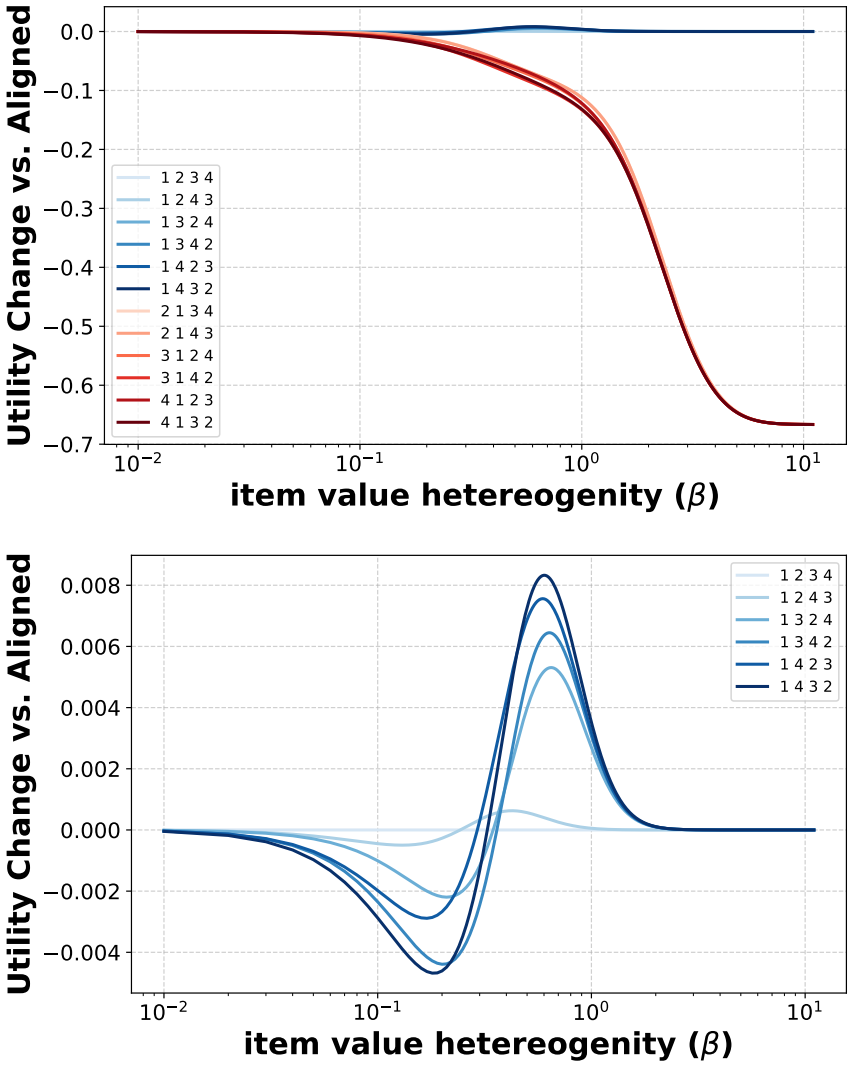


Figure 6: Comparison of human's expected utility differences under RUM.

Takeaway 3. *In the Mallows model and RUM, misalignment on relatively small-valued items still helps the human.*

9.2 Performance of MIP

We evaluate the performance of the MIP by varying both the number of items m and the number of types of humans n. In the first setting, we consider two humans with ground-truth rankings $\rho_1^* = (x_1, x_2, x_3, x_4, \ldots, x_m)$ and $\rho_2^* = (x_4, x_2, x_3, x_1, \ldots, x_m)$, where only x_1 and x_4 are swapped, and utilities $(v_{i,j})_{j=1}^m = 4, 3, 2, 1, 0 \ldots, 0$. We vary m from 10 to 70 and test $k = 2, 4, 6, 8$. For each test, we vary the frequencies of the first type of human from 0 to 1 and take the average of the running time. In the second setting, humans are drawn from a Mallows distribution with $n = t!$ possible types, $t = 1, \ldots, 6$. We fix $m = 20$, and test $k = 2, 4, 6, 8$. As shown in Figure 7, all instances terminate within 175 seconds, with the slowest cases taking 40.46 and 169.92 seconds, respectively. The running time grows polynomially in m and linearly in n.

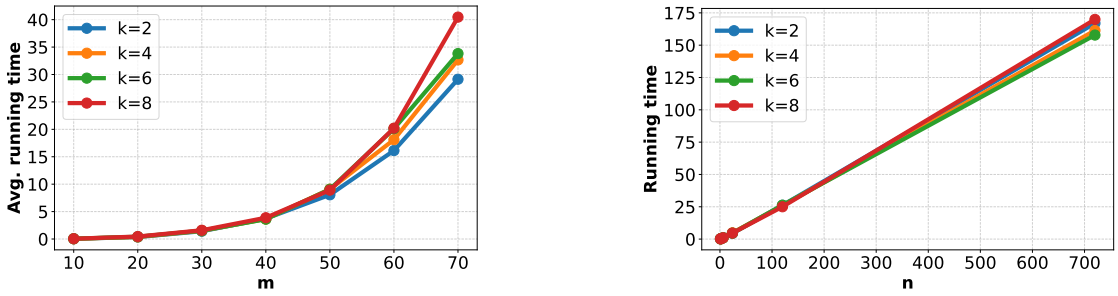


Figure 7: Running time of MIP (varying m and n)

Takeaway 4. *The welfare-maximizing algorithm can be found efficiently in practice.*

9.3 Tension between Welfare Objectives

The part mainly discusses the effect of enforcing uplift on social welfare. We consider $m = 6$ items and the humans' preferences differ only in the top three items, which follow a Mallows distribution with parameter γ, reflecting the balance of the populations. When γ is large, most humans have preferences aligned with $(x_1, x_2, \ldots, x_t)$. When γ is small, the humans' preferences become more dispersed. Each human's utilities are set as $(v_{i,j})_{j=1}^6 = 1, 1, 0.5, 0.2, 0, 0$. All humans share the same accuracy parameter ϕ_h, which we vary from 0 to 3 in increments of 0.3. Figure 8 plots the social welfare of the welfare-maximizing algorithms with and without the uplift constraint under different values of γ. The expected social welfare under the uplift constraint is consistently lower than that without it. When human preferences become more concentrated ($\gamma = 3$), the optimal solutions start to diverge. The largest welfare gap occurs at $\phi_h = 1$, where the expected social welfare without the uplift constraint is 0.989, compared to 0.954 with the constraint.

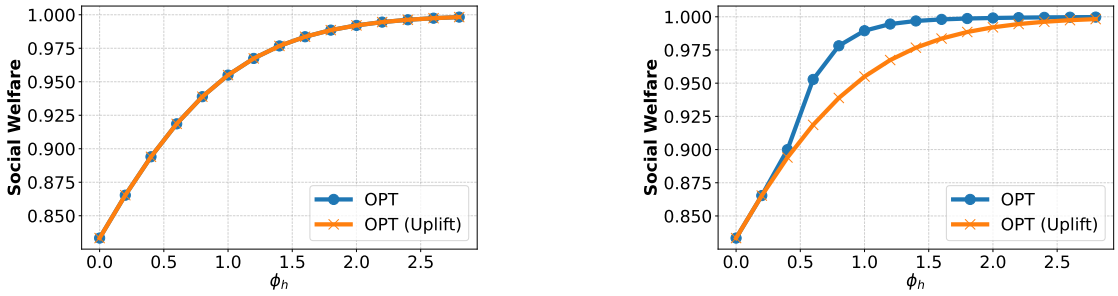


Figure 8: Max welfare: ±Uplift of $\gamma = 0.5$ (left) and 3 (right)

Takeaway 5. *Enforcing uplift reduces the optimal social welfare, especially when humans' distribution is imbalanced.*

Acknowledgements

The research of Bhaskar Ray Chaudhury and Jiaxin Song was supported by NSF CAREER Grant CCF-2441580. The research of Kate Donahue was supported by the MIT METEOR Fellowship. The research of Parnian Shahkar was supported by NSF Grant CCF-2454115. We thank Kira Goldner (Boston University), Michael Kearns (University of Pennsylvania), Ariel Procaccia (Harvard University), Aaron Roth (University of Pennsylvania), Jann Spiess (Stanford), Kostas Kollias (Google Research), Manish Raghavan (MIT), Biaoshuai Tao (SJTU), Ming Yin (Purdue), as well as attendees at the Themis Science Talk at Amazon, the Human-AI workshop at EC'25, the Human-AI Complementarity for Decision Making Workshop at CMU, the AI and Science Group at Meta, and the Boston Economics and Computing Hub Workshops, for their valuable discussions and insightful suggestions. We also thank the anonymous reviewers for their constructive feedback.

References

Parand A Alamdari, Soroush Ebadian, and Ariel D Procaccia. Policy aggregation. *Advances in Neural Information Processing Systems*, 37:68308–68329, 2024. 5

David H Ackley, Geoffrey E Hinton, and Terrence J Sejnowski. A learning algorithm for boltzmann machines. *Cognitive science*, 9(1):147–169, 1985. 3

Eugene Agichtein, Eric Brill, and Susan Dumais. Improving web search ranking by incorporating user behavior information. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 19–26, 2006. 14

Marina Agranov and Pietro Ortoleva. Stochastic choice and preferences for randomization. *Journal of Political Economy*, 125(1):40–68, 2017. 1, 14

Rohan Alur, Loren Laine, Darrick Li, Manish Raghavan, Devavrat Shah, and Dennis Shung. Auditing for human expertise. *Advances in Neural Information Processing Systems*, 36:79439–79468, 2023. 5

Rohan Alur, Manish Raghavan, and Devavrat Shah. Human expertise in algorithmic prediction. *Advances in Neural Information Processing Systems*, 37:138088–138129, 2024. 5

Anastasios Angelopoulos, Stephen Bates, Jitendra Malik, and Michael I Jordan. Uncertainty sets for image classifiers using conformal prediction. *arXiv preprint arXiv:2009.14193*, 2020. 5

Anastasios N Angelopoulos, Stephen Bates, et al. Conformal prediction: A gentle introduction. *Foundations and Trends® in Machine Learning*, 16(4):494–591, 2023. 5

Adrian Arnaiz-Rodriguez, Nina Corvelo Benz, Suhas Thejaswi, Nuria Oliver, and Manuel Gomez-Rodriguez. Towards human-ai complementarity in matching tasks. *arXiv preprint arXiv:2508.13285*, 2025. 5

Pranjal Awasthi, Avrim Blum, Or Sheffet, and Aravindan Vijayaraghavan. Learning mixtures of ranking models. *Advances in Neural Information Processing Systems*, 27, 2014. 6, 23, 33

Varun Babbar, Umang Bhatt, and Adrian Weller. On the utility of prediction sets in human-ai teams, 2022. 5

Gagan Bansal, Besmira Nushi, Ece Kamar, Eric Horvitz, and Daniel S Weld. Is the most accurate ai the best teammate? optimizing ai for teamwork. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 11405–11414, 2021a. 5

Gagan Bansal, Tongshuang Wu, Joyce Zhou, Raymond Fok, Besmira Nushi, Ece Kamar, Marco Tulio Ribeiro, and Daniel Weld. Does the whole exceed its parts? the effect of ai explanations on complementary team performance. In *Proceedings of the 2021 CHI conference on human factors in computing systems*, pages 1–16, 2021b. 4, 5, 10

Niclas Boehmer, Robert Bredereck, Piotr Faliszewski, and Rolf Niedermeier. A quantitative and qualitative analysis of the robustness of (real-world) election winners. In *Proceedings of the 2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, pages 1–10, 2022. 5

Niclas Boehmer, Piotr Faliszewski, and Sonja Kraiczy. Properties of the mallows model depending on the number of alternatives: a warning for an experimentalist. In *International Conference on Machine Learning*, pages 2689–2711. PMLR, 2023. 23, 25

Ralph Allan Bradley and Milton E Terry. Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, 39(3/4):324–345, 1952. 6

William Brown and Arpit Agarwal. Diversified recommendations for agents with adaptive preferences. *Advances in Neural Information Processing Systems*, 35:26066–26077, 2022. 5

William Brown and Arpit Agarwal. Online recommendations for agents with discounted adaptive preferences. In *International Conference on Algorithmic Learning Theory*, pages 244–281. PMLR, 2024. 5

Ioannis Caragiannis, Ariel D. Procaccia, and Nisarg Shah. When do noisy votes reveal the truth? *ACM Trans. Econ. Comput.*, 4(3), March 2016. ISSN 2167-8375. doi: 10.1145/2892565. URL https://doi.org/10.1145/2892565. 5

Lawrence Chan, Dylan Hadfield-Menell, Siddhartha Srinivasa, and Anca Dragan. The assistive multi-armed bandit. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 354–363. IEEE, 2019. 5

Daiwei Chen, Yi Chen, Aniket Rege, and Ramya Korlakai Vinayak. Pal: Pluralistic alignment framework for learning from heterogeneous preferences. *arXiv preprint arXiv:2406.08469*, 2024. 5

Rex Chen, Ruiyi Wang, Norman Sadeh, and Fei Fang. Missing pieces: How do designs that expose uncertainty longitudinally impact trust in ai decision aids? an in situ study of gig drivers. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*, pages 790–816, 2025. 5

Natalie Collina, Surbhi Goel, Varun Gupta, and Aaron Roth. Tractable agreement protocols. *arXiv preprint arXiv:2411.19791*, 2024. 5

Natalie Collina, Ira Globus-Harris, Surbhi Goel, Varun Gupta, Aaron Roth, and Mirah Shi. Collaborative prediction: Tractable information aggregation via agreement. *arXiv preprint arXiv:2504.06075*, 2025a. 5, 14

Natalie Collina, Surbhi Goel, Aaron Roth, Emily Ryu, and Mirah Shi. Emergent alignment via competition. *arXiv preprint arXiv:2509.15090*, 2025b. 5

Vincent Conitzer, Rachel Freedman, Jobst Heitzig, Wesley H Holliday, Bob M Jacobs, Nathan Lambert, Milan Mossé, Eric Pacuit, Stuart Russell, Hailey Schoelkopf, et al. Social choice for ai alignment: Dealing with diverse human feedback. *arXiv preprint arXiv:2404.10271*, 2024. 5

Nina Corvelo Benz and Manuel Rodriguez. Human-aligned calibration for ai-assisted decision making. *Advances in Neural Information Processing Systems*, 36:14609–14636, 2023. 5

Nina Laura Corvelo Benz and Manuel Gomez Rodriguez. Human-alignment influences the utility of ai-assisted decision making. *arXiv preprint arXiv:2501.14035*, 2025. 5

Bo Cowgill and Megan T Stevenson. Algorithmic social engineering. In *AEA Papers and Proceedings*, volume 110, pages 96–100, 2020. 5

Nick Craswell, Onno Zoeter, Michael Taylor, and Bill Ramsey. An experimental comparison of click position-bias models. In *Proceedings of the 2008 international conference on web search and data mining*, pages 87–94, 2008. 14

Jessica Dai and Eve Fleisig. Mapping social choice theory to rlhf. *arXiv preprint arXiv:2404.13038*, 2024. 5

Giovanni De Toni, Nastaran Okati, Suhas Thejaswi, Eleni Straitouri, and Manuel Gomez-Rodriguez. Towards human-ai complementarity with predictions sets. *arXiv preprint arXiv:2405.17544*, 2024. 5

Antoine Désir, Vineet Goyal, Srikanth Jagabathula, and Danny Segev. Assortment optimization under the mallows model. *Advances in Neural Information Processing Systems*, 29, 2016. 3, 32

Kate Donahue, Alexandra Chouldechova, and Krishnaram Kenthapadi. Human-algorithm collaboration: Achieving complementarity and avoiding unfairness. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, pages 1639–1656, 2022. 5

Kate Donahue, Sreenivas Gollapudi, and Kostas Kollias. When are two lists better than one?: Benefits and harms in joint decision-making, 2024. URL https://arxiv.org/abs/2308.11721. 3, 5, 10, 14, 23, 25, 26

Matteo Fontana, Gianluca Zeni, and Simone Vantini. Conformal prediction: a unified review of theory and new challenges. *Bernoulli*, 29(1):1–23, 2023. 5

Luise Ge, Daniel Halpern, Evi Micha, Ariel D Procaccia, Itai Shapira, Yevgeniy Vorobeychik, and Junlin Wu. Axioms for ai alignment from human feedback. *arXiv preprint arXiv:2405.14758*, 2024. 5

Paul Gölz, Nika Haghtalab, and Kunhe Yang. Distortion of ai alignment: Does preference optimization optimize for preferences? *arXiv preprint arXiv:2505.23749*, 2025. 5

Catalina Gomez, Sue Min Cho, Shichang Ke, Chien-Ming Huang, and Mathias Unberath. Human-ai collaboration is not very collaborative yet: a taxonomy of interaction patterns in ai-assisted decision making from a systematic review. *Frontiers in Computer Science*, 6:1521066, 2025. 5

Ben Green and Yiling Chen. The principles and limits of algorithm-in-the-loop decision making. *Proceedings of the ACM on Human-Computer Interaction*, 2019. 5

Sophie Greenwood, Karen Levy, Solon Barocas, Jon Kleinberg, and Hoda Heidari. Designing algorithmic delegates. In *NeurIPS 2024 Workshop on Behavioral Machine Learning*, 2024. 5

Ziyang Guo, Yifan Wu, Jason Hartline, and Jessica Hullman. The value of information in human-ai decision-making. *arXiv preprint arXiv:2502.06152*, 2025. 5

John D Hey and Chris Orme. Investigating generalizations of expected utility theory using experimental data. *Econometrica: Journal of the Econometric Society*, pages 1291–1326, 1994. 1, 14

Jessica Hullman, Yifan Wu, Dawei Xie, Ziyang Guo, and Andrew Gelman. Conformal prediction and human decision making. *arXiv preprint arXiv:2503.11709*, 2025. 5

Lujain Ibrahim, Katherine M Collins, Sunnie SY Kim, Anka Reuel, Max Lamparth, Kevin Feng, Lama Ahmad, Prajna Soni, Alia El Kattan, Merlin Stein, et al. Measuring and mitigating overreliance is necessary for building human-compatible ai. *arXiv preprint arXiv:2509.08010*, 2025. 5

Shomik Jain, Kathleen Creel, and Ashia Wilson. Scarce resource allocations that rely on machine learning should be randomized. *arXiv preprint arXiv:2404.08592*, 2024. 11, 14

Thorsten Joachims, Laura Granka, Bing Pan, Helene Hembrooke, Filip Radlinski, and Geri Gay. Evaluating the accuracy of implicit feedback from clicks and query reformulations in web search. *ACM Transactions on Information Systems (TOIS)*, 25(2):7–es, 2007. 14

Toshihiro Kamishima. Nantonac collaborative filtering: recommendation based on order responses. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 583–588, 2003. 12

Gerard Jounghyun Kim. *Human-computer interaction: fundamentals and practice*. CRC press, 2015. 5

Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. *Research methods in human-computer interaction*. Morgan Kaufmann, 2017. 5

Zhuoyan Li, Zhuoran Lu, and Ming Yin. Decoding ai's nudge: A unified framework to predict human behavior in ai-assisted decision making. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 10083–10091, 2024. 5

Allen Liu and Ankur Moitra. Efficiently learning mixtures of mallows models. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 627–638. IEEE, 2018. 6

R Duncan Luce et al. *Individual choice behavior*, volume 4. Wiley New York, 1959. 1, 4, 6, 14

I. Scott MacKenzie. *Human-computer interaction: An empirical research perspective*. Elsevier, 2024. 5

David Madras, Toni Pitassi, and Richard Zemel. Predict responsibly: Improving fairness and accuracy by learning to defer. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018. URL https://proceedings.neurips.cc/paper/2018/file/09d37c08f7b129e96277388757530c72-Paper.pdf. 5

Colin L Mallows. Non-null ranking models. i. *Biometrika*, 44(1/2):114–130, 1957. 4, 25

Bryce McLaughlin and Jann Spiess. Designing algorithmic recommendations to achieve human-ai complementarity. *arXiv preprint arXiv:2405.01484*, 2024. 14

Kanad Pardeshi, Itai Shapira, Ariel D Procaccia, and Aarti Singh. Learning social welfare functions. *Advances in Neural Information Processing Systems*, 37:41733–41766, 2024. 5

Kenny Peng, Nikhil Garg, and Jon Kleinberg. A no free lunch theorem for human-ai collaboration. *arXiv preprint arXiv:2411.15230*, 2024. 5

Robin L Plackett. The analysis of permutations. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 24 (2):193–202, 1975. 1, 4, 6, 14

Jenny Preece, Yvonne Rogers, Helen Sharp, David Benyon, Simon Holland, and Tom Carey. *Human-computer interaction*. Addison-Wesley Longman Ltd., 1994. 5

Karthik Raman and Thorsten Joachims. Methods for ordinal peer grading. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1037–1046, 2014. 6

Charvi Rastogi, Liu Leqi, Kenneth Holstein, and Hoda Heidari. A unifying framework for combining complementary strengths of humans and ml toward better predictive decision-making. *arXiv preprint arXiv:2204.10806*, 2022. 5

Ali Shirali, Arash Nasr-Esfahany, Abdullah Alomar, Parsa Mirtaheri, Rediet Abebe, and Ariel Procaccia. Direct alignment with heterogeneous preferences. *arXiv preprint arXiv:2502.16320*, 2025. 5

Ashudeep Singh and Thorsten Joachims. Fairness of exposure in rankings. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 2219–2228, 2018. 11, 14

Anand Siththaranjan, Cassidy Laidlaw, and Dylan Hadfield-Menell. Distributional preference learning: Understanding and accounting for hidden context in rlhf. *arXiv preprint arXiv:2312.08358*, 2023. 5

Taylor Sorensen, Jared Moore, Jillian Fisher, Mitchell Gordon, Niloofar Mireshghallah, Christopher Michael Rytting, Andre Ye, Liwei Jiang, Ximing Lu, Nouha Dziri, et al. A roadmap to pluralistic alignment. *arXiv preprint arXiv:2402.05070*, 2024. 5

Mark Steyvers, Heliodoro Tejeda, Gavin Kerrigan, and Padhraic Smyth. Bayesian modeling of human–ai complementarity. *Proceedings of the National Academy of Sciences*, 119(11):e2111547119, 2022. 5

Eleni Straitouri and Manuel Gomez Rodriguez. Designing decision support systems using counterfactual prediction sets, 2023. 5

Eleni Straitouri, Lequn Wang, Nastaran Okati, and Manuel Gomez Rodriguez. Provably improving expert predictions with conformal prediction, 2022. 5

Louis L Thurstone. A law of comparative judgment. *Psychological review*, 34(4):273, 1927. 4

Michelle Vaccaro, Abdullah Almaatouq, and Thomas Malone. When combinations of humans and ai are useful: A systematic review and meta-analysis. *arXiv preprint arXiv:2405.06087*, 2024. 5

Jingyan Wang and Ashwin Pananjady. Modeling and correcting bias in sequential evaluation. *arXiv preprint arXiv:2205.01607*, 2022. 14

Lequn Wang, Thorsten Joachims, and Manuel Gomez Rodriguez. Improving screening processes via calibrated subset selection. In *International Conference on Machine Learning*, pages 22702–22726. PMLR, 2022. 5

Dongping Zhang, Angelos Chatzimparmpas, Negar Kamali, and Jessica Hullman. Evaluating the utility of conformal prediction sets for ai-advised image labeling. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pages 1–19, 2024. 5

A Properties of Noisy Permutation Model

A.1 Basic Properties

Let $\mathfrak{S}(M)$ be the set of all rankings of M. Denote by $x_i \succ_\pi S$ for a set of items S if x_i is before any item $x_j \in S \backslash \{x_i\}$. A swap $(x_i x_j)$ is *valid* with respect to π and a ground-truth ranking π^* if π and π^* differ on the relative ranking of x_i and x_j. By the definition of inversion-monotonicity, if $(x_i x_j)$ is a valid swap with respect to any given ranking π and a ground-truth ranking π^*, then $\pi \circ (x_i, x_j)$ has a higher probability than π. Moreover, according to [Donahue et al., 2024, Lemma 1], the number of inversions reduces by at least one after the swap.

Lemma 9. *Let $\mathcal{D}$ be an inversion monotonic distribution with ground-truth ranking π^*. The random permutation π is sampled from $\mathcal{D}$. For any subset $\mathcal{S}$ of k items, the probability of $\mathcal{S}$ being the first k items of π decreases if substituting $x_i \in \mathcal{S}$ with another item $x_j \notin \mathcal{S}$ with $x_i \succ_{\pi^*} x_j$.*

Proof. Let $\mathcal{S}' = \mathcal{S} \setminus \{x_i\} \cup \{x_j\}$. Define two set of permutations $\mathfrak{S}_\mathcal{S}$ and $\mathfrak{S}_{\mathcal{S}'}$ as the set of permutations that place $\mathcal{S}$ and $\mathcal{S}'$ at the first k locations, respectively. Then we define a bijective mapping from a permutation $\pi \in \mathfrak{S}_\mathcal{S}$ to a permutation $\pi' \in \mathfrak{S}_{\mathcal{S}'}$ by swapping the items x_i and x_j. By inverse-monotonicity, the probability of π' is (weakly) higher than that of π. Therefore, we can conclude that the probability of $\mathcal{S}'$ being the first k of π is greater than that of $\mathcal{S}$. □

Lemma 10. *Let $\mathcal{D}$ be an inversion monotonic distribution with ground-truth ranking π^*. A permutation π is drawn from the distribution $\mathcal{D}$. Let $\mathcal{S}$ be a subset of k items such that $x_\ell \in \mathcal{S}$, $x_r \notin \mathcal{S}$, and $x_\ell \succ_{\pi^*} x_r$. Define $\mathcal{S}' = (\mathcal{S} \backslash x_\ell) \cup x_r$. Then, for any item $x_i \in \mathcal{S} \setminus \{x_\ell\}$, the probability that x_i is ranked first among $\mathcal{S}$ in π is less than the probability that it is ranked first among $\mathcal{S}'$ in π:*

$$\mathbb{P}[x_i \succ_\pi \mathcal{S}] \leq \mathbb{P}[x_i \succ_\pi \mathcal{S}'] .$$

The inequality becomes tight only when x_r and x_ℓ are totally indistinguishable under $\mathcal{D}$.

Proof. Let $\mathfrak{S}_R = \{\pi : x_i \text{ is before } \mathcal{S}' \text{ in } \pi\}$, and $\mathfrak{S}_L = \{\pi : x_i \text{ is before } \mathcal{S} \text{ in } \pi\}$. To show the inequality, it suffices to prove $\mathbb{P}[\pi \in \mathfrak{S}_R \setminus \mathfrak{S}_L] > \mathbb{P}[\pi \in \mathfrak{S}_L \setminus \mathfrak{S}_R]$. By definitions of $\mathfrak{S}_L$ and $\mathfrak{S}_R$, $\mathfrak{S}_R \setminus \mathfrak{S}_L$ contains the permutations where x_ℓ is placed before x_i and x_i is placed before other items of $\mathcal{S}'$ and x_r. Similarly, $\mathfrak{S}_L \setminus \mathfrak{S}_R$ consists of the permutations where x_r is placed before x_i and x_i is placed before other items of $\mathcal{S}$ (including x_ℓ). Define a mapping from $\mathfrak{S}_L \setminus \mathfrak{S}_R$ to $\mathfrak{S}_R \setminus \mathfrak{S}_R$ by swapping x_ℓ and x_r, which is a valid mapping by definition. As the distribution is inversion monotonic, the probability of the new permutation is greater than the original one, which implies that $\mathbb{P}[\pi \in \mathfrak{S}_R \setminus \mathfrak{S}_L] > \mathbb{P}[\pi \in \mathfrak{S}_L \setminus \mathfrak{S}_R]$ and concludes the lemma. □

A.2 Properties of Mallows Model

In Table 2, we provide an overview of the properties of the Mallows model used in our proofs, where $\pi \sim \mathcal{D}(\pi^*, \phi)$ and $Z_m(\phi) = \sum_{i=1}^m \exp(-\phi(i-1))$. Without loss of generality, we assume $\pi^* = (x_1, \ldots, x_m)$ in this subsection.

Event	Probability	Source
x_i ranked the first in π	$\frac{\exp(-\phi\cdot(i-1))}{Z_m(\phi)}$	[Awasthi et al., 2014]
x_1 ranked at the i-th location in π	$\frac{\exp(-\phi\cdot(i-1))}{Z_m(\phi)}$	[Awasthi et al., 2014]
x_1 ranked at the first k locations in π	$\frac{Z_k(\phi)}{Z_m(\phi)}$	[Awasthi et al., 2014]
x_i ranked before x_j with $i < j$ and $k = j - i + 1$ in π	$\frac{k}{1-\exp(-\phi\cdot k)} - \frac{k-1}{1-\exp(-\phi\cdot(k-1))}$	[Boehmer et al., 2023]
a set of items S being the first k items of π	closed-formed	Lemma 11
the first item of S in π is the best among S in π^*	$\geq \frac{1}{Z_k(\phi)}$ with $k = \lvert S \rvert$	Lemma 15
the first item of S in π is within the top s-th among S in π^*	$\geq \frac{Z_s(\phi)}{Z_k(\phi)}$ with $k = \lvert S \rvert$	Lemma 16

Table 2: Properties of the Mallows model.

A.2.1 Item Location

We first show that, for every set of $\mathcal{S}$ with k items, the probability of $\mathcal{S}$ being the first k items of π can have a closed-form expression in the Mallows model using the insertion-based sampling. Consider a fixed order $\pi_S = (x_{j(1)}, \ldots, x_{j(k)})$ of the k items. We insert the k items in order π_S. At the time of inserting the i-th one, the probability of placing $x_{j(i)}$ at the i-th location conditioned on the previous insertion, depends on the relative location of $x_{j(1)}$ among $\mathcal{S}$ and the inserted items. Consequently, it implies the probability of the prefix is solely determined by the number of inversions in π_S. By summing it up for all π_S, we get a closed-form probability of $\mathcal{S}$ being the first k, which only depends on the sum of indices of items of $\mathcal{S}$. The formal theorem statement and proof are presented as follows:

Lemma 11. *In a Mallows distribution $\pi \sim \mathcal{D}(\pi^*, \phi)$, given a subset S of items of size k, the probability of set S of being the first k items of π is given by*

$$\mathbb{P}[\pi[:k] = S] = \exp\left(-\phi \cdot \left(\sum\nolimits_{x_j \in S} j - \frac{k(k+1)}{2}\right)\right) \cdot \prod_{i=1}^{k} \frac{Z_i(\phi)}{Z_{m-i+1}(\phi)}\,.$$

Proof. Fix an order π_S of S. Below, we give the probability of π_S being the prefix of π. Denote the k items by $x_{j(1)}, \ldots, x_{j(k)}$ according order π_S. We insert the k items in order. At time of inserting the i-th item $x_{j(i)}$, consider the probability of locating $x_{j(i)}$ at the i-th place in π conditioned on $x_{j(1)}, \ldots, x_{j(i-1)}$ being located as the first $i-1$ items of π. Observe that the number of inserted items that are placed before $x_{j(i)}$ in the ground truth ranking π^* is equal to $|\{i' : j(i') < j(i) \wedge i' < i\}|$. Denote that number by $A(i)$. The current item $x_{j(i)}$ is the $(j(i) - A(i))$-th item among the set of remaining items $M \setminus \{x_{j(t)}\}_{t=1}^{i-1}$. Then, according to the property of the Mallows model, the probability of $x_{j(i)}$ being placed in i-th place in π conditioned on the past insertion is given by

$$\mathbb{P}\left[\pi(i) = x_{j(i)} \mid \wedge_{t=1}^{i-1}\pi(t) = x_{j(t)}\right] = \frac{1}{Z_{m-i+1}(\phi)} \cdot \exp\left(-\phi \cdot (j(i) - A(i) - 1)\right)\,. \tag{2}$$

Based on it, the probability of π_S being the prefix of π is equal to the product of these conditional probabilities, which is given by

$$\begin{aligned}
\mathbb{P}\left[\wedge_{i=1}^{k}\pi(i) = x_{j(i)}\right] &= \prod_{i=1}^{k} \Pr\left[\pi(i) = x_{j(i)} \mid \wedge_{t=1}^{i-1}\pi(t) = x_{j(t)}\right] && \text{(By the chain rule)} \\
&= \prod_{i=1}^{k} \frac{\exp\left(-\phi \cdot (j(i) - A(i) - 1)\right)}{Z_{m-i+1}(\phi)} && \text{(By Equation (2))}
\end{aligned}$$

By direct calculation, we can see that

$$\mathbb{P}\left[\wedge_{i=1}^{k}\pi(i) = x_{j(i)}\right] = \exp\left(-\phi \cdot \sum_{i=1}^{k}(j(i) - i + i - 1 - A(i))\right) \cdot \prod_{i=1}^{k} \frac{1}{Z_{m-i+1}(\phi)}\,.$$

Notice that the term $i - 1 - A(i)$ is essentially the number of items that are placed before $x_{j(i)}$ in π but ranked after $x_{j(i)}$ in the ground-truth ranking π^*. Hence, each of them forms an inversion with $x_{j(i)}$. Thus, the sum of $i - 1 - A(i)$ over i equals to the number of inversions of π_S. Therefore, we have

$$\mathbb{P}\left[\wedge_{i=1}^{k}\pi(i) = x_{j(i)}\right] = \exp\Big(-\phi \cdot \Big(\sum_{x_j \in S} j - \frac{k(k+1)}{2}\Big)\Big) \cdot \exp(-\phi \cdot \mathrm{inv}(\pi_S)) \cdot \prod_{i=1}^{k} \frac{1}{Z_{m-i+1}(\phi)}\,. \tag{3}$$

By summing the above equation over all the permutations of items in S, we can obtain that

$$\mathbb{P}[\pi[:k] = S] = \sum_{\pi_S \in \mathfrak{S}(S)} \mathbb{P}\left[\wedge_{i=1}^{k}\pi(i) = x_{j(i)}\right]\,.$$

Then applying Equation (3), since the sum of $\exp(-\phi \cdot \mathrm{inv}(\pi_S))$ over all permutations of S is equal to $\prod_{i=1}^{k} Z_i(\phi)$ [Mallows, 1957], we then have

$$\mathbb{P}\left[\pi(:k) = S\right] = \exp\left(-\phi \cdot \left(\sum_{x_j \in S} j - \frac{k(k+1)}{2}\right)\right) \cdot \frac{\prod_{i=1}^{k} Z_i(\phi)}{\prod_{i=1}^{k} Z_{m-i+1}(\phi)} \,. \qquad \square$$

A.2.2 Itemwise Comparison

Lemma 12 (Proposition 3.8 of [Boehmer et al., 2023]). *In a Mallows distribution $\pi \sim \mathcal{D}(\pi^*, \phi)$, for any two distinct items $x_i, x_j \in M$ with $i < j$ and $k = j - i + 1$, the probability that x_i is before x_j in π is given by*

$$\Pr\left[x_i \succ_\pi x_j\right] = \frac{k}{1 - \exp(-\phi \cdot k)} - \frac{k-1}{1 - \exp(-\phi \cdot (k-1))} \,.$$

Lemma 13 (Monotonicity of Pairwise Comparison). *In a Mallows distribution $\pi \sim \mathcal{D}(\pi^*, \phi)$, the probability of the pairwise comparison between items x_i and x_j with $j > i$ increases with the index gap $j - i$. Particularly, $\mathbb{P}[x_i \succ_\pi x_j] \leq \mathbb{P}[x_{i'} \succ_\pi x_{j'}]$ if $j' - i' \geq j - i$.*

Proof. We first prove it holds for the case of $i = i'$. Hence, by assumption, $j' \geq j$. For any ranking π that ranks x_i before x_j while ranking x_i after $x_{j'}$, we map it to a new one by swapping x_j and $x_{j'}$. By [Donahue et al., 2024, Lemma 1], the number of inversions reduces by at least one, and the probability of the new ranking is weakly larger than the original one. Therefore, it implies that the probability of x_i being before x_j is smaller. By Lemma 12, the probability of pairwise comparison only depends on the difference of the two indices. Therefore, the lemma holds for every i, j, i', and j'. $\square$

Lemma 14. *For any n and a constant $\epsilon < 1/2$, there exists $\phi > 0$ and m such that in Mallows distribution $\pi \sim \mathcal{D}(\pi^*, \phi)$ with ground-truth ranking $\pi^* = (x_1, \ldots, x_m)$, for any two items x_i and x_j, the following holds*

$$\mathbb{P}[x_i \succ_\pi x_j] \in \begin{cases} (0, \frac{1}{2} + \epsilon], & \text{if } j - i + 1 \leq n \\ [1 - \epsilon, 1), & \text{if } j - i + 1 \geq m - n \end{cases}$$

Proof. When $j - i + 1 \leq n$, we first show an upper bound of the probability of x_i being before x_j. Similarly, we define a bijective mapping from the rankings where x_i is before x_j to rankings where x_i is after x_i by swapping the location of x_i and x_j. Since $j - 1 + 1 \leq n$, the number of inversions will increase by at most n. Hence, $\mathbb{P}[x_i \succ_\pi x_j] \leq \exp(\phi \cdot n)$, which immediately implies that, $\mathbb{P}[x_i \succ_\pi x_j] \leq \frac{\exp(\phi \cdot n)}{1 + \exp(\phi \cdot n)}$. Thus, by direct calculation, it suffices to set $\phi = \frac{1}{n} \cdot \log(\frac{1+2\epsilon}{1-2\epsilon})$ for the first condition.

Next, we find the number m satisfying the second condition. By Lemma 13, for any fixed ϕ, the probability increases as k. Hence, it suffices to show it holds for $m - n$. Recall that, by Lemma 12,

$$\mathbb{P}[x_i \succ_\pi x_j] = \frac{k}{1 - \exp(-\phi \cdot k)} - \frac{k-1}{1 - \exp(-\phi \cdot (k-1))} \,.$$

Direct calculation gives that

$$\begin{aligned} \mathbb{P}[x_i \succ_\pi x_j] &= \frac{1}{1 - \exp(-\phi \cdot (k-1))} \cdot \left(1 - \frac{\exp(-\phi \cdot (k-1)) - \exp(-\phi \cdot k)}{1 - \exp(-\phi \cdot k)} k\right) \\ &\geq \frac{1}{1 - \exp(-\phi \cdot (k-1))} \cdot (1 - \exp(-\phi \cdot (k-1)) \cdot k) \,. \end{aligned}$$

To make sure the probability is greater than $1 - \epsilon$, it suffices to set $\exp(-\phi \cdot (k-1)) \leq \epsilon / k$, which can be satisfied when $k / \log k \geq 1/\phi$. As $\log k \leq \sqrt{k}$, it suffices to set $m - n \geq k \geq 1/\phi^2$. Therefore, we set $\phi = \frac{1}{n} \cdot \log(\frac{1+2\epsilon}{1-2\epsilon})$ and $m = 1/\phi^2 + n$. $\square$

Lemma 15. *In a mallows distribution $\pi \sim \mathcal{D}(\pi^*, \phi)$, given a subset of items $\mathcal{S} = \{x_{i(1)}, \ldots, x_{i(k)}\}$ with $i(1) < \ldots < i(k)$, the probability of $x_{i(1)}$ being placed before all other items of $\mathcal{S}$ in π satisfies*

$$\mathop{\mathbb{P}}_{\pi \sim \mathcal{D}(\pi^*, \phi)} \left[x_{i(1)} \succ_\pi \mathcal{S} \right] \geq \frac{1}{Z_k(\phi)} .$$

Proof. We partition the set of permutations into k sets based on the relative position of $x_{i(1)}$. Let $\mathfrak{S}_t$ be the set of permutations where $x_{i(1)}$ is placed at the t-th relative position among all items of $\mathcal{S}$. For any $\pi \in \mathfrak{S}_{t+1}$ with $t \leq k-1$, we map it to a permutation $\pi \in \mathfrak{S}_t$ by swapping $x_{i(1)}$ and the item that is placed at the t-th position among all items of $\mathcal{S}$. It is a valid swapping with respect to π and π^* by definition. By [Donahue et al., 2024, Lemma 1], the probability of π' is at least $\exp(\phi)$ times the probability of π. By summing the inequality up over all permutations of $\mathfrak{S}_{t+1}$, we can conclude that $\mathbb{P}[\pi \in \mathfrak{S}_t] \geq \exp(\phi) \cdot \mathbb{P}[\pi \in \mathfrak{S}_{t+1}]$. Hence, by multiplying these inequalities, we can have

$$\mathbb{P}[\pi \in \mathfrak{S}_1] \geq \exp(\phi \cdot (t-1)) \cdot \mathbb{P}[\pi \in \mathfrak{S}_t] .$$

Therefore, by summing it over all t, we have

$$\mathbb{P}[\pi \in \mathfrak{S}_1] \geq \frac{1}{\sum_{t=1}^{k} \exp(\phi \cdot (t-1))} \cdot \sum_{t=1}^{k} \mathbb{P}[\pi \in \mathfrak{S}_t] = \frac{1}{Z_k(\phi)}. \qquad \square$$

Lemma 16. *In a mallows distribution $\pi \sim \mathcal{D}(\pi^*, \phi)$ with $\pi^* = (x_1, \ldots, x_m)$, given a subset of items $\mathcal{S} = \{x_{i(1)}, \ldots, x_{i(k)}\}$ with $i(1) < \ldots < i(k)$, for any s such that $1 \leq s \leq k$, the probability of one of $x_{i(1)}, \ldots, x_{i(s)}$ being the first among all items of $\mathcal{S}$ is at least*

$$\sum_{j=1}^{s} \mathbb{P} \left[x_{i(j)} \succ_\pi \mathcal{S} \right] \geq \frac{Z_s(\phi)}{Z_k(\phi)} .$$

Proof. Similar to the above proof, let $\mathfrak{S}_j$ be the set of permutations where $x_{i(j)}$ is the first one among all the items in $\mathcal{S}$. Then the left-hand side can be rewritten as $\sum_{j=1}^{s} \mathbb{P}[\pi \in \mathfrak{S}_j]$. Denote each term by $f(j)$ for $j = 1, \ldots, s$. Using the argument of valid-swapping-based mapping, we have $f(i) \geq \exp(\phi) \cdot f(i+1)$. Meanwhile, as $\sum_{i=1}^{k} f(i) = 1$, we have

$$\begin{aligned}
&\sum_{j=1}^{s} f(j) - \frac{Z_s(\phi)}{Z_k(\phi)} = \sum_{j=1}^{s} f(j) - \frac{Z_s(\phi)}{Z_k(\phi)} \cdot \sum_{j=1}^{k} f(j) \\
&\propto \sum_{j=1}^{s} f(j) \cdot (Z_k(\phi) - Z_s(\phi)) - \sum_{j=s+1}^{k} f(j) \cdot Z_s(\phi) && \text{(By multiplying } Z_k(\phi)) \\
&= \sum_{j=1}^{s} f(j) \cdot \exp(-\phi \cdot s) \cdot Z_{k-s}(\phi) - \sum_{j=s+1}^{k} f(j) \cdot Z_s(\phi) && (Z_k(\phi) - Z_s(\phi) = \tfrac{Z_{k-s}(\phi)}{\exp(\phi \cdot s)})
\end{aligned}$$

Since $Z_{k-s}(\phi)$ can be unfolded as $Z_{k-s}(\phi) = \sum_{t=1}^{k-s} \exp(-\phi(t-1))$, by changing the order of summation, we can obtain that

$$\begin{aligned}
\sum_{j=1}^{s} f(j) - \frac{Z_s(\phi)}{Z_k(\phi)} &\propto \sum_{t=1}^{k-s} \exp(-\phi \cdot (t-1)) \cdot \sum_{j=1}^{s} f(j) \cdot \exp(-\phi \cdot s) - \sum_{j=1}^{k} f(j) \cdot Z_s(\phi) \\
&= \sum_{t=s+1}^{k} \exp(-\phi \cdot (t-1)) \cdot \sum_{j=1}^{s} f(j) - \sum_{j=s+1}^{k} f(j) \cdot Z_s(\phi) \\
&= \sum_{t=s+1}^{k} \sum_{j=1}^{s} (f(j) \cdot \exp(-\phi \cdot (t-1))) - \sum_{j=s+1}^{k} f(j) \cdot Z_s(\phi) \\
&\geq \sum_{t=s+1}^{k} \sum_{j=1}^{s} f(t) \cdot \exp(-\phi \cdot (j-1)) - \sum_{j=s+1}^{k} f(j) \cdot Z_s(\phi) && (f(i) \geq f(i+1))
\end{aligned}$$

$$= \sum_{t=s+1}^{k} f(t) \cdot Z_s(\phi) - \sum_{j=s+1}^{k} f(j) \cdot Z_s(\phi) = 0,$$

which concludes the proof. □

B Omitted Details of Section 4

B.1 Omitted Proofs of Theorem 1

Lemma 1. *For any item that is not i, the probability of picking that item is higher with the misaligned algorithm (and the probability of picking item i is lower). That is, for any $r \in [m]$ with $r \neq i$, $\mathbb{P}[x_C^1 = x_r] \leq \mathbb{P}[x_C^2 = x_r]$, and $\mathbb{P}[x_C^1 = x_i] \geq \mathbb{P}[x_C^2 = x_i]$*[6].

Proof. Let π_2^* be the ground-truth ranking of A_2, which swaps the ranking x_i and x_j compared to A_1's ground-truth ranking π_1^*. Let $\mathcal{D}_a^1$ and $\mathcal{D}_a^2$ be the distributions with ground-truth ranking π_1^* and π_2^* respectively. In addition, denote by x_C^1 and x_C^2 the two items picked by the human working with Algorithm 1 and Algorithm 2, respectively.

Case I: Picking item other than x_i and x_j. We first consider the change in the probability of picking some item other than x_i or x_j. Without loss of generality, denote by x_r the considered item. Denote by x_C^1 and x_C^2, respectively, the items picked by the human when collaborating with algorithm 1 and algorithm 2. By the way of human-algorithm interaction, x_r is chosen by a human only if it is included in the presented list $\mathcal{S}$ (i.e., $\pi[:k] = \mathcal{S}$) and the human ranks it before any other item of $\mathcal{S}$ (i.e., $x_r \succ_\rho \mathcal{S}$). Thus, the probabilities of x_r being picked by the human are given by

$$\mathbb{P}[x_C^1 = x_r] = \sum_{\mathcal{S}: x_r \in \mathcal{S}} \mathbb{P}\left[\pi_1[:k] = \mathcal{S}\right] \times \mathbb{P}\left[x_r \succ_\rho \mathcal{S}\right], \tag{Prob 1}$$

$$\mathbb{P}[x_C^2 = x_r] = \sum_{\mathcal{S}: x_r \in \mathcal{S}} \mathbb{P}\left[\pi_2[:k] = \mathcal{S}\right] \times \mathbb{P}\left[x_r \succ_\rho \mathcal{S}\right], \tag{Prob 2}$$

where $\pi_1 \sim \mathcal{D}_a^1$, $\pi_2 \sim \mathcal{D}_a^2$, and $\rho \sim \mathcal{D}_h$.

Next, we conduct a term-by-term comparison of Prob 1 and Prob 2 with respect to each $\mathcal{S}$. We start with two simple cases: $\mathcal{S}$ contains both x_i and x_j; $\mathcal{S}$ contains neither x_i nor x_j. For either of the two cases, consider the permutation π_1 whose first k items are $\mathcal{S}$. Swap x_i and x_j in π_1 and get another permutation π_2, which shares the same set of the first k items. As the two distributions only differ in the relative rankings of x_i and x_j in their ground-truth rankings, the probability of π_1 occurring in $\mathcal{D}_a^1$ is the same as π_2 occurring in $\mathcal{D}_a^2$. Therefore, the two summations are equal in these terms.

Next, we compare the two probabilities by pairing the remaining summation terms corresponding to $\mathcal{S}$ that contains only one of x_i and x_j. Denote by $\mathcal{S}^i$ a set $\mathcal{S}$ containing x_i without x_j. We pair $\mathcal{S}^i$ with another subset by substituting item x_i with x_j. Denote the new set by $\mathcal{S}^j$. Next, we show that the following inequality holds for every constructed pair $(\mathcal{S}^i, \mathcal{S}^j)$,

$$\sum_{\mathcal{S} \in (\mathcal{S}^i, \mathcal{S}^j)} \mathbb{P}\left[\pi_1[:k] = \mathcal{S}\right] \mathop{\mathbb{P}}_{\rho \sim \mathcal{D}_h}\left[x_r \succ_\rho \mathcal{S}\right] \leq \sum_{\mathcal{S} \in (\mathcal{S}^i, \mathcal{S}^j)} \mathbb{P}\left[\pi_2[:k] = \mathcal{S}\right] \mathop{\mathbb{P}}_{\rho \sim \mathcal{D}_h}\left[x_r \succ_\rho \mathcal{S}\right]. \tag{InEq (1)}$$

As x_i is better than x_j in human's ground-truth ranking ρ^*, it is easier for the human to rank x_r before x_j, as described in Lemma 10. Formally,

$$\mathbb{P}[x_r \succ_\rho \mathcal{S}^j] \geq \mathbb{P}[x_r \succ_\rho \mathcal{S}^i].$$

Also, since the two distributions $\mathcal{D}_a^1$ and $\mathcal{D}_a^2$ are isomorphic, differing only by a swap between x_i and x_j in ground-truth rankings, by relabeling the two items, we have

$$\mathop{\mathbb{P}}_{\pi_1 \sim \mathcal{D}_a^1}\left[\pi_1[:k] = \mathcal{S}^i\right] = \mathop{\mathbb{P}}_{\pi_2 \sim \mathcal{D}_a^2}\left[\pi_2[:k] = \mathcal{S}^j\right],$$

[6]Note that the two inequalities become tight only when the two items are completely indistinguishable to the human, i.e., swapping them does not change the probability of any ranking under the human's distribution. The following content considers the case where this condition does not hold.

$$\mathop{\mathbb{P}}_{\pi_1\sim\mathcal{D}_a^1}\left[\pi_1[:k]=\mathcal{S}^j\right]=\mathop{\mathbb{P}}_{\pi_2\sim\mathcal{D}_a^2}\left[\pi_2[:k]=\mathcal{S}^i\right]\ .$$

Meanwhile, by Lemma 9, as x_i is placed before x_j in π_1^*, we have

$$\mathop{\mathbb{P}}_{\pi_1\sim\mathcal{D}_a^1}\left[\pi_1[:k]=\mathcal{S}^i\right]>\mathop{\mathbb{P}}_{\pi_1\sim\mathcal{D}_a^1}\left[\pi_1[:k]=\mathcal{S}^j\right]\ .$$

By applying the rearrangement inequality, then InEq (1) holds.

Case II: Picking x_i or x_j. We next consider the change of the probabilities of picking x_i or x_j. First, we show that the probability of picking x_i decreases after the algorithm places x_i after x_j. Similarly,

$$\mathbb{P}[x_C^1=x_i]=\sum_{\mathcal{S}:x_i\in\mathcal{S}}\mathbb{P}\left[\pi_1[:k]=\mathcal{S}\right]\cdot\mathop{\mathbb{P}}_{\rho\sim\mathcal{D}_h}\left[x_i\succ_\rho\mathcal{S}\right],$$
$$\mathbb{P}[x_C^2=x_i]=\sum_{\mathcal{S}:x_i\in\mathcal{S}}\mathbb{P}\left[\pi_2[:k]=\mathcal{S}\right]\cdot\mathop{\mathbb{P}}_{\rho\sim\mathcal{D}_h}\left[x_i\succ_\rho\mathcal{S}\right].$$

Next, we still compare the two probabilities by term. Consider the terms where the set $\mathcal{S}$ contains both x_i and x_j. The first terms $\mathbb{P}[\pi_1[:k]=\mathcal{S}]$ and $\mathbb{P}[\pi_2[:k]=\mathcal{S}]$ are the same by identical logic used above. Next, consider the other terms where $\mathcal{S}$ does not contain x_j. Let $\mathcal{S}'=\mathcal{S}\setminus\{x_i\}\cup\{x_j\}$. By a similar reason, $\mathbb{P}[\pi_2[:k]=\mathcal{S}]=\mathbb{P}[\pi_1[:k]=\mathcal{S}']$. In addition, as x_i is placed before x_j in π_1^*, then by Lemma 9, $\mathbb{P}[\pi_1[:k]=\mathcal{S}]>\mathbb{P}[\pi[:k]=\mathcal{S}']$, which further concludes the first inequality. The remaining comparison of the probabilities that the human selects x_j under the two algorithms follows by symmetry. □

Theorem 1. *The human gains **higher** utility with the misaligned algorithm when x_i and x_j are least valued, but **lower** utility when x_i is top valued.*

Proof. Consider the change in the probability of the human picking each item. By the above lemma, after swapping x_i to a later position in algorithm's ground-truth, only the probability of the human picking item x_i decreases, while the probability of picking any other item increases. Therefore, if both x_i and x_j are zero-valued to the human, then such a swap leads to an increase in the probability of picking any valuable item, which in turn increases human's utility. For the second bullet, if x_i is the most valuable item to the human, then the swap causes a loss in picking the most valuable item, which further leads to a decrease in human's utility. □

Theorem 2 (Best/worst strategy for top item recovery)**.** *In the top item recovery setting, the algorithm's ground-truth ranking π^* that maximizes human's expected utility is $\pi^*=(x_1,x_m,\ldots,x_2)$ while the one minimizing human's expected utility is $\pi^*=(x_2,\ldots,x_m,x_1)$.*

Proof. By the first bullet of Theorem 1, swapping any pair of zero-valued items will increase the human's utility. Therefore, for the algorithm that maximizes human's utility, all items other than x_1 should be ranked in reverse order. In addition, by the second bullet, ranking human's top item x_1 in any place other than the first position will decrease human's utility. Thus, the optimal algorithm should rank x_1 first. Using a symmetric argument, we can also prove that the ranking $(x_2,x_3,\ldots,x_m,x_1)$ yields the least benefit.

Notably, this result extends to the setting where the human has multiple top-valued items. Suppose the human assigns a positive value only to her top d items, which are equally valued. In this case, the algorithm's arrangement π^* that maximizes the human's expected utility places the top d items first, followed by the remaining items in reverse order; that is, $\pi^*=(x_1,\ldots,x_d,x_m,\ldots,x_{d+1})$. □

B.2 Extensions of Theorem 1

Lemma 2. *Sufficient condition for when collaboration with a misaligned algorithm is **harmful** for the human under the Mallows model with accuracy parameter ϕ_h is $v_1-v_i\le\frac{\exp(-\phi_h\Delta)}{1-\exp(-\phi_h\Delta)}\Delta$ where $\Delta=j-i$ and $k=2$.*

Lemma 4. *Sufficient condition for when collaboration with a misaligned algorithm is **harmful** for the human under the Plackett-Luce model is when $v_0-v_j\le 1.27\beta$.*

Proof. According to the computation of Lemma 1, for every $r \neq i, j$, we have

$$\mathbb{P}[x_C^2 = x_r] - \mathbb{P}[x_C^1 = x_r] = (\mathbb{P}[\pi_1[:2] = \{x_i, x_r\}] - \mathbb{P}[\pi_1[:2] = \{x_j, x_r\}]) \cdot (\mathbb{P}[x_r \succ_\rho x_j] - \mathbb{P}[x_r \succ_\rho x_i]) \geq 0$$

Meanwhile, we have

$$\mathbb{P}[x_C^2 = x_j] - \mathbb{P}[x_C^1 = x_j] = \sum_{r \neq i,j} (\mathbb{P}[\pi_1[:2] = \{x_i, x_r\}] - \mathbb{P}[\pi_1[:2] = \{x_j, x_r\}]) \cdot \mathbb{P}[x_j \succ_\rho x_r] \geq 0$$

$$\mathbb{P}[x_C^2 = x_i] - \mathbb{P}[x_C^1 = x_i] = \sum_{r \neq i,j} (\mathbb{P}[\pi_1[:2] = \{x_i, x_r\}] - \mathbb{P}[\pi_1[:2] = \{x_j, x_r\}]) \cdot \mathbb{P}[x_i \succ_\rho x_r] \leq 0$$

Let $\psi(i, j, r)$ be $\psi(i, j, r) = \mathbb{P}[\pi_1[:2] = \{x_i, x_r\}] - \mathbb{P}[\pi_1[:2] = \{x_j, x_r\}]$. Therefore, the expected change in human's utility is given by

$$\begin{aligned}
&\sum_{r \neq i,j} v_r \cdot (\mathbb{P}[x_C^2 = x_r] - \mathbb{P}[x_C^1 = x_r]) + v_j \cdot (\mathbb{P}[x_C^2 = x_j] - \mathbb{P}[x_C^1 = x_j]) + v_i \cdot (\mathbb{P}[x_C^2 = x_i] - \mathbb{P}[x_C^1 = x_i]) \\
&= \sum_{r \neq i,j} v_r \cdot \psi(i, j, r) \cdot (\mathbb{P}[x_r \succ_\rho x_j] - \mathbb{P}[x_r \succ_\rho x_i]) + \sum_{r \neq i,j} v_j \cdot \psi(i, j, r) \cdot \mathbb{P}[x_j \succ_\rho x_r] - \sum_{r \neq i,j} v_i \cdot \psi(i, j, r) \cdot \mathbb{P}[x_i \succ_\rho x_r] \\
&= \sum_{r \neq i,j} v_r \cdot \psi(i, j, r) \cdot (\mathbb{P}[x_i \succ_\rho x_r] - \mathbb{P}[x_j \succ_\rho x_r]) + \sum_{r \neq i,j} v_j \cdot \psi(i, j, r) \cdot \mathbb{P}[x_j \succ_\rho x_r] - \sum_{r \neq i,j} v_i \cdot \psi(i, j, r) \cdot \mathbb{P}[x_i \succ_\rho x_r] \\
&= \sum_{r \neq i,j} \psi(i, j, r) \cdot ((v_r - v_i) \cdot (\mathbb{P}[x_i \succ_\rho x_r]) - (v_r - v_j) \cdot (\mathbb{P}[x_j \succ_\rho x_r]))
\end{aligned}$$

Next, we show that the above summation is always negative when the preconditions of the two models are satisfied. Notice that, when $i \leq r \leq j$, by assumption, we have $v_i \geq v_r \geq v_j$. Hence, the inner term is always negative (as $v_i \neq v_j$). In addition, when $r \geq j$, by Lemma 10, we know $\mathbb{P}[x_i \succ_\rho x_r] \geq \mathbb{P}[x_j \succ_\rho x_r]$. Meanwhile, we can observe that $|x_r - x_i| \geq |x_r - x_j|$. Hence, the inner term is still negative. It suffices to prove that every term with $r < i$ is negative.

(a) Plackett-Luce model. Let $\delta = v_r - v_i$ and $\Delta = v_i - v_j$. By the definition of the Plackett-Luce model, the inner term can be rewritten as

$$(v_r - v_i) \cdot \frac{\exp(v_i/\beta)}{\exp(v_r/\beta) + \exp(v_i/\beta)} - (v_r - v_j) \cdot \frac{\exp(v_j/\beta)}{\exp(v_r/\beta) + \exp(v_j/\beta)} = \frac{\delta}{\exp(\delta/\beta) + 1} - \frac{\delta + \Delta}{\exp((\delta + \Delta)/\beta) + 1}.$$

Let function $f(x) = \frac{x}{\exp(x/\beta)+1}$. Since $f'(x) = \frac{(1-x/\beta)\exp(x/\beta)+1}{(\exp(x/\beta)+1)^2}$, then $f(x)$ is increasing when $x \leq 1.278$. Therefore, the above term is always negative since $\delta + \Delta = v_r - v_j \leq v_0 - v_j \leq 1.27\beta$.

(b) Mallows model. Since swapping item x_i and x_j at most increases the number of inversion by $|j - i|$, then $\mathbb{P}[x_j \succ_\rho x_r] \geq \exp(-\phi_h \cdot (i - j)) \cdot \mathbb{P}[x_i \succ_\rho x_r]$. Since $v_r \leq v_0 \leq \frac{\exp(\phi_h \cdot (i-j))}{\exp(\phi_h \cdot (i-j)) - 1} v_i - \frac{1}{\exp(\phi_h \cdot (i-j)) - 1} v_j$, we have $v_r - v_i \leq \exp(-\phi_h \cdot (i - j)) \cdot (v_r - v_j)$, which means that the inner term is always non-positive and concludes the proof. □

Lemma 3. *Sufficient condition for when collaboration with a misaligned algorithm is* ***helpful*** *for the human under the Mallows model is when some $i' \in [i-1]$ satisfies*

$$\frac{v_{i'}}{v_i} \geq \frac{\sum_{r \neq i,j} \psi(i, j, r)}{\sum_{r=1}^{i'} \psi(i, j, r) \exp(-\phi_h \cdot (j - r + 1))} \frac{1}{1 - \exp(-\phi_h \cdot \Delta)},$$

where $\Delta = j - i + 1$ and $\psi(i, j, r) = \mathbb{P}[\pi_1[:2] = \{x_i, x_r\}] - \mathbb{P}[\pi_1[:2] = \{x_j, x_r\}]$ is always nonnegative.

Lemma 5. *Sufficient condition for when collaboration with a misaligned algorithm is* ***helpful*** *under the Plackett-Luce model for the human is when some $i' \in [i-1]$ satisfies*

$$\frac{v_{i'}}{v_i} \geq \frac{\sum_{r \neq i,j} \frac{\psi(i,j,r)}{\exp((v_r - v_i)/\beta) + 1}}{\sum_{r=1}^{i'} \frac{\psi(i,j,r)}{\exp((v_r - v_i)/\beta) + 1}} \cdot \frac{2}{1 - \exp(-\Delta/\beta)},$$

where $\Delta = v_i - v_j$ and $\psi(i, j, r) = \mathbb{P}[\pi_1[:2] = \{x_i, x_r\}] - \mathbb{P}[\pi_1[:2] = \{x_j, x_r\}] \geq 0$.

Proof. Let i' be the index satisfying the given condition in each of the two models, respectively. Then the expected change is at least

$$\sum_{r\neq i,j} v_r\cdot\psi(i,j,r)\cdot(\mathbb{P}[x_r\succ_\rho x_j]-\mathbb{P}[x_r\succ_\rho x_i])+\sum_{r\neq i,j} v_j\cdot\psi(i,j,r)\cdot\mathbb{P}[x_j\succ_\rho x_r]-\sum_{r\neq i,j} v_i\cdot\psi(i,j,r)\cdot\mathbb{P}[x_i\succ_\rho x_r]$$
$$>\sum_{r=1}^{i'} v_r\cdot\psi(i,j,r)\cdot(\mathbb{P}[x_r\succ_\rho x_j]-\mathbb{P}[x_r\succ_\rho x_i])-\sum_{r\neq i,j} v_i\cdot\psi(i,j,r)\cdot\mathbb{P}[x_i\succ_\rho x_r]$$

(a) Plackett-Luce model. When the human's ranking satisfies the Plackett-Luce model, then the above value is at least

$$\sum_{r=1}^{i'} v_r\cdot\psi(i,j,r)\cdot\left(\frac{\exp(v_r/\beta)}{\exp(v_j/\beta)+\exp(v_r/\beta)}-\frac{\exp(v_r/\beta)}{\exp(v_i/\beta)+\exp(v_r/\beta)}\right)-\sum_{r\neq i,j} v_i\cdot\psi(i,j,r)\cdot\frac{\exp(v_i/\beta)}{\exp(v_i/\beta)+\exp(v_r/\beta)}$$
$$\geq\sum_{r=1}^{i'} v_r\cdot\psi(i,j,r)\cdot\frac{\exp(v_r/\beta)\cdot(\exp(v_i/\beta)-\exp(v_j/\beta))}{(\exp(v_i/\beta)+\exp(v_r/\beta))(\exp(v_j/\beta)+\exp(v_r/\beta))}-\sum_{r\neq i,j} v_i\cdot\psi(i,j,r)\cdot\frac{\exp(v_i/\beta)}{\exp(v_i/\beta)+\exp(v_r/\beta)}$$
$$\geq\sum_{r=1}^{i'} v_r\cdot\psi(i,j,r)\cdot\frac{\exp(v_r/\beta)\cdot(\exp(v_i/\beta)-\exp(v_j/\beta))}{(\exp(v_i/\beta)+\exp(v_r/\beta))^2}-\sum_{r\neq i,j} v_i\cdot\psi(i,j,r)\cdot\frac{\exp(v_i/\beta)}{\exp(v_i/\beta)+\exp(v_r/\beta)}$$
$$\geq v_{i'}\cdot\sum_{r=1}^{i'}\psi(i,j,r)\cdot\frac{\exp(v_r/\beta)\cdot(\exp(v_i/\beta)-\exp(v_j/\beta))}{(\exp(v_i/\beta)+\exp(v_r/\beta))^2}-v_i\cdot\sum_{r\neq i,j}\psi(i,j,r)\cdot\frac{\exp(v_i/\beta)}{\exp(v_i/\beta)+\exp(v_r/\beta)}$$
$$\geq v_{i'}\cdot\sum_{r=1}^{i'}\psi(i,j,r)\cdot\frac{\exp(v_i/\beta)-\exp(v_j/\beta)}{2(\exp(v_i/\beta)+\exp(v_r/\beta))}-v_i\cdot\sum_{r\neq i,j}\psi(i,j,r)\cdot\frac{\exp(v_i/\beta)}{\exp(v_i/\beta)+\exp(v_r/\beta)}$$
$$\geq 0\,.$$

(b) Mallows model. When the human's ranking satisfies Mallows model, we can notice that for any $r\leq i'$,

$$\begin{aligned}\mathbb{P}[x_r\succ_\rho x_j]-\mathbb{P}[x_r\succ_\rho x_i]&\geq\frac{1}{1-\exp(-\phi_h\cdot(r-j+1))}-\frac{1}{1-\exp(-\phi_h\cdot(r-i))}\\&\geq\frac{\exp(-\phi_h\cdot(r-j+1))-\exp(-\phi_h\cdot(r-i))}{(1-\exp(-\phi_h\cdot(r-i)))\cdot(1-\exp(-\phi_h\cdot(r-j+1)))}\\&\geq\exp(-\phi_h\cdot(r-j+1))\cdot(1-\exp(-\phi_h\cdot(i-j-1))\end{aligned}$$

Therefore, the expected change is at least

$$\sum_{r=1}^{i'} v_r\cdot\psi(i,j,r)\cdot\exp(-\phi_h\cdot(r-j+1))\cdot(1-\exp(-\phi_h\cdot(i-j-1))-\sum_{r\neq i,j} v_i\cdot\psi(i,j,r)\cdot\mathbb{P}[x_i\succ_\rho x_r]$$
$$\geq v_{i'}\cdot(1-\exp(-\phi_h\cdot(i-j-1))\cdot\sum_{r=1}^{i'}\psi(i,j,r)\cdot\exp(-\phi_h\cdot(r-j+1))-v_i\sum_{r\neq i,j}\psi(i,j,r)$$
$$\geq 0,\qquad\text{(By assumption of } v_{i'} \text{ and } v_i)$$

which concludes the lemma. □

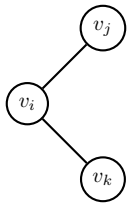
C Omitted Proofs in Section 5

C.1 NP-Hardness of Welfare-Maximization

Theorem 3 (Computational hardness)**.** *The expected social welfare is maximized when the algorithm's distribution is noiseless. However, it remains* NP*-hard to find π^* that maximizes the expected social welfare, even in the* top-item recovery *setting.*

Proof. As discussed, the optimal solution is noiseless. Next, we reduce from the Independent Set problem, where the input is a graph instance $G = (V, E)$ with n vertices and m edges, and an integer k. The output is **YES** if the size of the maximum independent set of graph G is at least k and **NO** otherwise. We construct the human-algorithm collaboration instance in the following way.

Fix $\epsilon > 0$ to be a sufficiently small constant such that $(1-\epsilon)^{k-1} > 1 - 1/(2k)$ and $\epsilon < 1/4$. Construct m items $x_1, \ldots, x_m$. Denote the n vertices as $v_1, \ldots, v_n$. For each $i \in [n]$, we construct a ground-truth ranking ρ_i^* as follows: place x_i as the first item in ρ_i^*. Iterate through all the other vertices from v_1 to v_n and insert x_j after x_i in ρ_i^* if vertex v_i is adjacent to vertex v_j in the input graph instance G. Otherwise, put v_j to the end of ρ_i^*. For the remaining items $x_{n+1}, \ldots, x_m$, put them in the middle of the inserted items. Set $p_i = 1/n$ for each $i \in [n]$ and $v_1 = 1$, $v_i = 0$ for $i \geq 2$.



Input graph G

$$\rho_j^*: \quad x_j \succ x_i \succ x_{n+1} \succ \cdots \succ x_m \succ x_j$$

$$\rho_i^*: \quad x_i \succ x_j \succ x_{n+1} \succ \cdots \succ x_m \succ x_k$$

$$\rho_k^*: \quad x_k \succ x_i \succ x_{n+1} \succ \cdots \succ x_m \succ x_j$$

Constructed rankings ρ_i^*, ρ_j^*, and ρ_k^*

We choose ϕ_h and m so that in a Mallows model with m items and accuracy parameter ϕ_h, given two items, the probability of ranking before the other is at most than $1/2 + \epsilon$ when the difference of the indices is smaller than n and is larger than $1 - \epsilon$ when it is larger than $m - n$, which can be achieved by Lemma 14.

Next, we first show that if the input graph instance is a **YES** instance, then the optimal expected social welfare is at least $(1-\epsilon)^k \cdot k/n$. Suppose the vertex index set of the maximum independent set of G is I with $|I| \geq k$. Put k of the independent set as the first k items of π^* and the other $m - k$ items in an arbitrary order. Hence, specifying $\mathcal{S} = \pi^*[:k]$, for every human $i \in I$, the probability of picking her best item is at least

$$\mathbb{P}[x_C^i = x_i \mid \mathcal{S}] = \mathbb{P}[x_i \succ_{\rho_i} \mathcal{S}] = \mathbb{P}\left[\wedge_{j \in I} (x_i \succ_{\rho_i} x_j)\right] \geq \prod_{j \in I} \mathbb{P}\left[x_i \succ_{\rho_i} x_j\right] \geq (1-\epsilon)^{k-1}.$$

Therefore, the expected utility of the human with ground-truth ρ_i^*, $\mathbb{E}[u(x_J \mid \rho_i^*)]$ is at least $(1-\epsilon)^{k-1}$. As the human has a probability of $p_i = 1/n$ of having ground-truth ρ_i^* for every $i \in I$, the expected social welfare is at least $(1-\epsilon)^{k-1} \cdot k/n$.

Next, we consider when the input instance is **NO**. Denote the k items of π^* by $\mathcal{S}$. As the size of $\mathcal{S}$ is k, there are $n - k$ humans whose best item is not picked. Hence, their utilities have a probability of at least $(1-\epsilon)$ being 0. In addition, as the input instance is **NO** instance, consider the vertices corresponding to $\mathcal{S}$. There must be two adjacent vertices included at the same time. Hence, there will be two humans (say i, j) whose best item is within the top n items of the other one's ground-truth ranking. By the assumption of ϕ_h, the two humans only have the probability of no more than $1/2 + \epsilon$ to pick their best item. Therefore, the expected social welfare is at most

$$\begin{aligned}
&\frac{1}{n} \cdot \left(\mathbb{E}[u_i(x_C^i)] + \mathbb{E}[u_j(x_C^j)] + \sum_{t \in I \setminus \{i,j\}} \mathbb{E}[u_t(x_C^t)] + \sum_{t \notin I} \mathbb{E}[u_t(x_C^t)] \right) \\
&\leq \frac{1}{n} \cdot \left(2 \cdot \left(\frac{1}{2} + \epsilon \right) + (k-2) \cdot 1 \right) \\
&< \frac{1}{n} \cdot (k - 1 + 2\epsilon) < \frac{1}{n} \cdot k \cdot (1-\epsilon)^{k-1}, \qquad \text{(as } (1-\epsilon)^{k-1} > 1 - 1/(2k) \text{ and } \epsilon < 1/4)
\end{aligned}$$

which is smaller than the expected social welfare under the **YES** instance. Therefore, the problem of finding a welfare-maximizing strategy is NP-hard for the algorithm. □

Observation 5. *We observe that for a fixed arrangement π^*, noiselessness is not always optimal in terms of the social welfare. Consider the following example. The algorithm's ground-truth is fixed as $\pi^* = (x_1, x_2, x_3)$, but items x_1 and x_2 only have positive value to a proportion of 1% of people.*

Algorithm 1: Dynamic programming for choice probability without prior knowledge

Function `ChoiceProb`$(\mathcal{S}, x_i)$**:**
  Input: a subset $\mathcal{S}$ of items $[m]$;
  Output: Probability of item x_i being the first among all items in $\mathcal{S}$;
  for $t = 1, \ldots, m$ **do**
    for *each* $1 \le i \le k-1$ *and* $s = 1, \ldots, t$ **do**
      $W(i, s, t) \leftarrow (1 - \gamma_{t,s}) \cdot W(i, s, t-1) + \mathbf{1}[x_t \notin \mathcal{S}] \cdot \gamma_{t,s-1} \cdot W(i, s-1, t-1)$;
    for *each* $s = 1, \ldots, t$ **do**
      $W(t, s, t) \leftarrow \mathbf{1}[x_t \in \mathcal{S}] \cdot p_{t,s} \cdot \left(\sum_{i \le t-1} \sum_{\ell=s}^{m} W(i, \ell, t-1) + \prod_{i=1}^{t-1} \mathbf{1}[x_i \notin \mathcal{S}] \right)$;
  return $\sum_{s=1}^{m} W(i, s, m)$;

If the algorithm is noiseless and only presents two items to the humans, it causes 99% of people to get no benefit, which results in worse social welfare than a noisy algorithm that would randomize over items that are presented.

C.2 Mixed Integer Linear Programming Formulation

Finally, we adapt the MIP of Désir et al. [2016] and formulate the problem of finding the welfare-maximizing strategy as an MIP. Désir et al. [2016] studies the *assortment optimization* problem under the Mallows model, where each item is associated with a nonnegative value r_i and the goal is to pick a subset of items $\mathcal{S}$ to maximize the weighted sum of choice probabilities. Formally, the problem can be expressed as follows:

$$\max_{\mathcal{S} \subseteq [m]} \left(\sum_{x_i \in \mathcal{S}} r_i \cdot \mathbb{P}[x_i \succ \mathcal{S}] \right).$$

Désir et al. [2016] shows the choice probability of x_i being the first among all items in $\mathcal{S}$ can be calculated by a dynamic programming algorithm. However, it is worth noting that the algorithm relies on prior knowledge that some item is known to be included in $\mathcal{S}$, which is further used as a "guess" of the MIP formulation. We next provide a dynamic programming of the choice probability with marginal change to the original algorithm of Désir et al. [2016] without any prior knowledge.

Let $W(i, s, t)$ be the probability of item x_i being chosen at position s after t steps of repeatedly insertion, where $i \in [m]$, $s \in [t]$, and $t \in [m]$. The following algorithm assumes the ground-truth ranking $\pi^* = (x_1, \ldots, x_m)$ without loss of generality. We use the same notations for constants $p_{t+1,s}$, $\gamma_{t+1,s}$ as in Désir et al. [2016]. Probability $p_{t+1,s}$ is the probability of inserting item x_{t+1} at position s into a partial permutation of $x_1, \ldots, x_t$. Also, $\gamma_{t,s} = \sum_{\ell=1}^{s} p_{t,\ell}$. Both sets of constants have closed-form expressions once the ground-truth ranking and the accuracy parameter are given.

Lemma 17. *Algorithm 7 computes the choice probability of item x_i being the first among $\mathcal{S}$.*

Proof. Compared to the original dynamic programming of Désir et al. [2016], we do not require a precondition that an item x_i needs to be included in $\mathcal{S}$. Hence, we update the state space starting from step $t = 1$ rather than $t = 2$. As a result, if $x_t \in \mathcal{S}$, at the t-th step, there are two possible cases for item x_t being selected at position s: 1) none of items $x_1, \ldots, x_{t-1}$ is included in $\mathcal{S}$; therefore, item x_t is selected at position s with probability $p_{t,s}$; 2) at least one of items $x_1, \ldots, x_{t-1}$ is included in $\mathcal{S}$; therefore, item x_t is chosen only if all items of $x_1, \ldots, x_{t-1}$ that are included in $\mathcal{S}$ are ranked after x_t and x_t is inserted at position s (the same argument of Désir et al. [2016]). Otherwise, if $x_t \notin \mathcal{S}$, the same argument of Désir et al. [2016] still applies for updating the probabilities of choosing other items. □

MIP Formulation. Based on the above dynamic programming, we first give a simplified MIP formulation of the assortment problem in Fig. 9. For notational simplicity, we reuse σ to denote the ground-truth ranking. Note that, we no longer force σ to be $\{x_1, \ldots, x_m\}$. Denote by $\sigma(i)$ the index of the i-th item. It is worth noting, the new MIP does not rely on the guess of knowing some item included in $\mathcal{S}$. Therefore, it is no longer needed to enumerate all the possible guesses, which further reduces the running time by a factor of $O(m)$.

Based on the MIP, we then formulate the problem of finding the welfare-maximizing algorithm as an MIP as follows. We replicate the constraint of choice probability n times – once for each of the n Mallows distributions corresponding to the n types of humans. Notably, as each type of human has a different ground-truth ranking, σ is set accordingly in each case. In addition, the objective function is changed accordingly by summing the utilities of all humans.

C.3 Uplift under Special Cases

Lemma 6. *Suppose human of type i only has positive values for the top T items, i.e., $v_{i,j} > 0$ for $j \le T$ and $v_{i,j} = 0$ for $j > T$. Then if $\pi^*(j) = \rho^*(j)$ for any $j \le T$, $\phi_a = \phi_h$, and $1 < k < m$, uplift is achieved.*

Proof. We first show the lemma holds for a special case: $\phi_a = \phi_h = \phi$. Unfold the expectation of the two sides of the inequality of uplift. The original inequality is equivalent to the following inequality

$$\sum_{j=1}^{m} v_j \cdot \mathbb{P}[x_C = x_j] > \sum_{j=1}^{m} v_j \cdot \mathbb{P}[x_H = x_j]\,. \tag{4}$$

To show the above inequality, we apply Karamata's inequality to prove the following proposition, which essentially states that, for any $i \in [T]$, the probability of the joint system picking an item of $x_1, \ldots, x_i$ is weakly larger than the probability of the human acting alone.

Proposition 1. *For any $i \in [T]$, we have $\sum_{j=1}^{i} \mathbb{P}[x_C = x_j] > \sum_{j=1}^{i} \mathbb{P}[x_H = x_j]$.*

Proof. According to the law of conditional probability, we can expand the above probability conditioned on the first k items of human's permutation as follows,

$$\sum_{j=1}^{i} \mathbb{P}[x_H = x_j] = \sum_{j=1}^{i} \sum_{\mathcal{S}:|\mathcal{S}|=k} \mathbb{P}[x_H = x_j \mid \rho[:k] = \mathcal{S}] \cdot \mathbb{P}[\rho[:k] = \mathcal{S}]$$

By Lemma 11, the probability of $\mathbb{P}[\rho[:k] = \mathcal{S}]$ only depends on ϕ and the sum of indices of items of $\mathcal{S}$. As ρ^* an π^* have the same set of top T items, we can see $\mathbb{P}[\rho[:k] = \mathcal{S}] = \mathbb{P}[\pi[:k] = \mathcal{S}]$ and the above sum can be rewritten as

$$\sum_{j=1}^{i} \mathbb{P}[x_H = x_j] = \sum_{S:|S|=k} \mathbb{P}[\pi[:k] = S] \cdot \sum_{j=1}^{i} \mathbb{P}[x_H = x_j \mid \rho[:k] = S]\,.$$

Fix a set S of size k. Notice that, when $x_j \notin S$, $\mathbb{P}[x_H = x_j \mid \rho[:k] = \mathcal{S}] = 0$. Next, we consider the sum of the conditional probabilities for $x_j \in S$. We first show that, conditioned on $\rho[:k] = \mathcal{S}$, the partial ordering of the first k items (denoted by $\rho_{|\mathcal{S}}$) forms a Mallows distribution of items of $\mathcal{S}$. Observe that the number of inversions between S and $M \setminus S$ for every permutation ρ such that $\rho[:k] = \mathcal{S}$ since the two sets of items are isolated. Hence, among all the permutations such that $\rho[:k] = \mathcal{S}$, the probability of each of them is proportional to $\exp(-\phi)$ raised to the power of the number of inversions inside $\mathcal{S}$, which implies that the partial ranking $\pi_{|S}$ forms a Mallows distribution of items of $\mathcal{S}$. Denote the s items by $\mathcal{S} = \{x_{i(1)}, \ldots, x_{i(k)}\}$ with $i(1) < \ldots < i(k)$. By the known property of Mallows model [Awasthi et al., 2014, Lemma 2.5], the probability of item $x_{i(j)}$ being the first is equal to $\exp(-\phi \cdot (j-1))/Z_k(\phi)$. Summing it up for j from 1 to i, we get the probability of x_H being one of x_1 to x_i is equal to $Z_s(\phi)/Z_k(\phi)$, where s is the size of the intersection of $\{x_1, \ldots, x_i\}$ and $\mathcal{S}$. Therefore, the probability of the human picking one of the first i items is equal to

$$\sum_{j=1}^{i} \mathbb{P}[x_H = x_j] = \sum_{S:|S|=k} \mathbb{P}[\pi[:k] = \mathcal{S}] \cdot \frac{Z_s(\phi)}{Z_k(\phi)}, \quad \text{with } s = |S \cap \{x_1, \ldots, x_i\}|\,.$$

On the other side, by the definition of the pick-and-choose, the joint system picks the human's favorite item among the presented k items of the algorithm. Hence, we have

$$\sum_{j=1}^{i} \mathbb{P}[x_C = x_j] = \sum_{\mathcal{S}:|\mathcal{S}|=k} \mathbb{P}[\pi[:k] = \mathcal{S}] \cdot \sum_{j=1}^{i} \mathbb{P}[x_j \succ_\rho \mathcal{S}]\,. \tag{5}$$

In Lemma 16, we show that $\sum_{j=1}^{i} \mathbb{P}[x_j \succ_\rho S] \geq Z_s(\phi)/Z_k(\phi)$. It immediately implies that $\sum_{j=1}^{i} \mathbb{P}[x_C = x_j] \geq \sum_{j=1}^{i} \mathbb{P}[x_H = x_j]$ and concludes the proof. □

Using Karamata's inequality and Proposition 1, we conclude that the theorem holds. □

C.4 Small Noiseness Helps Uplift

Lemma 7. *There exist settings where uplift can occur at lower accuracy (higher noise), but fails at higher accuracy.*

Proof. Suppose there are three types of humans, and an algorithm is presenting $k = 2$ items to the human. The ground-truth rankings of the three types of humans are respectively $\rho_1^* = (x_1, x_2, \cdots, x_m)$, $\rho_2^* = (x_2, x_3, \cdots, x_m)$, and $\rho_3^* = (x_3, x_1, \cdots, x_m)$. Consider the top-item recovery setting: $v_1 = 1$ and $v_i = 0$ for $i \geq 2$. The accuracy parameter of the human ϕ_h is assumed to be sufficiently small such that every human is almost equally likely to pick any item by themselves.

If the algorithm sets the response to be noiseless, then the presented two items will be deterministic, which means only two of $\{x_1, x_2, x_3\}$ can be presented. As a result, at least one of the three types of humans is unable to pick their best item and receives no utility from the collaboration.

Nevertheless, once we add some noise to the response of the algorithm, uplift becomes possible. Consider setting the ground-truth ranking of the algorithm as $\pi^* = (x_1, x_2, x_3, \ldots, x_m)$ and the accuracy ϕ_a such that $\mathbb{P}[\pi(1) = x_1] = 1/2, \mathbb{P}[\pi(1) = x_2] = 1/4$, and $\mathbb{P}[\pi(3) = x_3] = 1/8$. For every type of human, if her favorite item is included in the presented two items, she has a probability of at least $1/2$ to choose it among the two items. Therefore, the probability of choosing the best item reaches at least $1/8 \cdot 1/2 = 1/16$, which beats the value of acting alone. □

C.5 NP-Hardness of Deciding Existence of Uplifting Strategy

Theorem 4. *It is NP-hard to determine whether there exists a strategy satisfying uplift (π^*, ϕ_a). Further, it is NP-complete when k is given as a constant (which means whether a given strategy (π^*, ϕ_a) achieves uplift can be verified in polynomial time).*

Proof. We reduce from the Vertex Cover problem. The input of the Vertex Cover problem is a graph instance $G = (V, E)$ along with a number k. The output is **YES** if there exists a vertex cover (containing at least one vertex of every edge) of size at most k and **NO** otherwise.

We construct a human-AI collaboration model as follows. By the properties of the Mallows model, for a fixed accuracy ϕ_h and ground-truth ranking ρ^*, in a Mallows distribution $\rho \sim \mathcal{D}(\rho^*, \phi_h)$, we have

$$\mathbb{P}_{\rho \sim \mathcal{D}(\rho^*, \phi_h)}[\rho(1) \in \{\rho^*(1), \rho^*(2)\}] = \frac{Z_2(\phi_h)}{Z_m(\phi_h)} > \frac{Z_2(\phi_h)}{1/(1 - \exp(-\phi))},$$
$$\mathbb{P}_{\rho \sim \mathcal{D}(\rho^*, \phi_h)}[\rho(1) \in \{\rho^*(1), \rho^*(2)\}] = \frac{Z_2(\phi_h)}{Z_m(\phi_h)} < \frac{Z_2(\phi_h)}{1 + \exp(-\phi_h) + \exp(-\phi_h \cdot 2)}.$$

Then set ϵ_1 as a sufficiently small constants such that $1 - \epsilon_1^{1/k} > 2\epsilon_1^{1/(2k)}$ and $\epsilon_1 < 1/4$, and the accuracy parameter ϕ_h such that $1 - \epsilon_1 < 1 - \exp(-\phi_h \cdot 2)$. Let ϵ_2 be defined as follows

$$\epsilon_2 = 1 - \frac{1 + \exp(-\phi_h)}{1 + \exp(-\phi_h) + \exp(-\phi_h \cdot 2)}.$$

Further, set ϵ_3 be a sufficiently small constant such that $(1 - \epsilon_3)^n \geq 1 - \epsilon_2$. Then we fix ϕ_h and apply Lemma 14 to find B such that the probability of pairwise comparison is no less than $1 - \epsilon_3$ once the index difference is no less than B, and also set B larger than n.

Construct $n + B$ items in total: n vertex items x_v for any vertex $v \in V$ and B dummy items y_i for $i \in [B]$. Also, we create one type of human for every edge $e = (u, v) \in E$ with ground-truth ranking

$$\rho_e^* = (x_u, x_v, y_1, \cdots, y_B, \cdots), \quad \text{for any } (u, v) \in E$$

which places items x_u and x_v first, followed immediately by all the dummy items, and finally the remaining vertex items. Each human only has positive values for the first two items $v_1 = v_2 = 1$ while $v_j = 0$ for $j > 2$.

We next prove the mapping between the two instances. When the input vertex cover is a **YES** instance, there exists a vertex cover $U \subseteq V$ that covers every edge. Consider the strategy that puts U first in π^* and set $\phi_a = +\infty$. For every human with ground-truth ranking ρ_e^*, if both x_u and x_v are included in U, then the probability of choosing either x_u or x_v strictly increases. Otherwise, if only x_u or x_v is included in U, as all other vertex items in U are placed to the end of ρ_e^*, the probability of pairwise comparison is at least $1 - \epsilon_2$. By the setting of ϵ_3, the probability of ranking x_u or x_v before other items in U is at least $(1 - \epsilon_3)^n \geq 1 - \epsilon_2$. Thus, the human still has a higher probability of choosing her favorite item and receives a higher expected utility. Hence, $(\pi^*, +\infty)$ achieves uplift and the human-AI model is also a **YES** instance.

We next consider when the input graph is a **NO** instance of the vertex cover problem. We prove it by contradiction. Suppose an uplift strategy exists. We first claim that the algorithm should be "very random". Consider the probability of the algorithm presenting the first k items of its ground-truth ranking. By Lemma 11, the probability of that event is greater than $1/\Pi_{i=1}^{k} Z_m(\phi_a)$. As the input instance is **NO** instance, there must exist an agent whose top two items are not included in the first k items of π^*. As a result, one human receives a utility of zero in this case. Since the utility of the human is at least $1 - \epsilon_1$ when acting alone, then

$$(1 - \exp(-\phi_a))^k < \frac{1}{\Pi_{i=1}^{k} Z_m(\phi_a)} < \epsilon_1 \implies \exp(-\phi_a) > 1 - \epsilon_1^{1/k}.$$

On the other hand, we prove that, to make sure the algorithm uplifts every human, the algorithm should also be "very robust" since it should make sure at least one of the favorite items of every human included in $\mathcal{S}$ with a high probability. By Lemma 9, the probability of an item ranked in the first k positions reaches maximum when the item is the first one in the ground-truth ranking. Hence, the maximum probability is $Z_k(\phi_a)/Z_{n+B}(\phi_a)$. Therefore,

$$\begin{aligned} 1 - \epsilon_1 &< \mathbb{P}[\mathcal{S} \cap \{x_u, x_v\} \neq \emptyset] \leq 1 - \mathbb{P}[\mathcal{S} \cap \{x_u\} = \emptyset] \cdot \mathbb{P}[\mathcal{S} \cap \{x_v\} = \emptyset] \\ &< 1 - \left(1 - \frac{Z_k(\phi_a)}{Z_{n+B}(\phi_a)}\right)^2 < 1 - \left(1 - \frac{1}{1 + \exp(-\phi_a)^k}\right)^2. \qquad \text{(as } B > n\text{)} \end{aligned}$$

Direct calculation implies that

$$1 - \frac{1}{1 + \exp(-\phi_a)^k} < \sqrt{\epsilon_1} \implies \exp(-\phi_a)^k < \frac{\sqrt{\epsilon_1}}{1 - \sqrt{\epsilon_1}} \implies \exp(-\phi_a) < \frac{\epsilon_1^{1/(2k)}}{(1 - \sqrt{\epsilon_1})^{1/k}} < 2\epsilon_1^{1/(2k)},$$

which contradicts the previous inequality $\exp(-\phi_a) > 1 - \epsilon_1^{1/k}$. Therefore, the human-AI model is also a **NO**-instance and it concludes the reduction.

NP-**membership.** Next, we provide a polynomial-time method to validate whether an algorithm with a fixed ground-truth ranking π^* and accuracy parameter ϕ_a achieves uplift and complementarity when the number of presented items k is constant-bounded. By the definition of the pick-and-choose, a human picks a particular item x_i only if the algorithm includes it in $\mathcal{S}$ and the human ranks it before other items of $\mathcal{S}$. Notice that the number of possible outcomes of $\mathcal{S}$ is equal to n^k.

For every set of $\mathcal{S}$ with k items, as shown in Section A.2.1, there is a closed-form expression of the probability of $\mathcal{S}$ being the first k items of π, which can be computed in $O(m)$ time. Moreover, Algorithm 7 can calculate the probability of an item x_i being first among a shortlist of items $\mathcal{S}$ in the Mallows model. Combining them, we can identify the exact utility of every human receiving from the collaboration, which further determines whether uplift is achieved or not. □

Corollary 1. *The problem of finding the welfare-maximizing strategy remains* NP*-hard under the top item recovery setting.*

Lemma 8. *In the top item recovery setting, when the ground-truth rankings of humans satisfy $|\mathcal{M}_0| \leq m - 1$ where $\mathcal{M}_0$ is the set of distinct items that some human has positive value for, then always presenting $\mathcal{M}_0$ to the human achieves uplift and also maximizes the expected utilitarian social welfare among all the noiseless algorithms that achieve uplift.*

Proof. Uplift stems from the algorithm aiding people in eliminating valueless items. When a human acts alone, she must choose her most valuable item (the "superstar") from the entire item set. As every valuable item remains

included in $\mathcal{M}_0$, each human can choose her favorite item among a shorter list, which leads to a higher probability than acting alone. Meanwhile, to maximize the maximin share, the algorithm should include every item of $\mathcal{M}_0$ in the presented item list $\mathcal{S}$. Otherwise, there must exist one type of human with no chance to pick her best item, and the minimum utility will be equal to 0, which is smaller than always presenting $\mathcal{M}_0$. Furthermore, for every item that is not favored by any human, erasing it from $\mathcal{S}$ only improves the probability of every human picking her best item, as it causes less misleading. Thus, the strategy also maximizes the expected utilitarian social welfare among all noiseless strategies. □

D Table of Section 6

Expanding on the numerical study presented in Section 6, the following table lists the most common sushi preference rankings among the 120 observed rankings. Each row reports a specific ranking, the number of participants (out of 5000) who expressed that preference, and the cumulative fraction of the total population accounted for up to that row. Notably, the top 33 most frequent rankings together represent the preferences of half of the participants.

Table 3: Population distribution and cumulative fractions for sushi rankings.

Population	Ranking	Cumulative Fraction of Total Population
126	[4, 5, 2, 1, 3]	0.0252
120	[4, 5, 1, 2, 3]	0.0492
119	[4, 5, 2, 3, 1]	0.0730
119	[2, 1, 3, 5, 4]	0.0968
119	[2, 3, 1, 5, 4]	0.1206
115	[1, 2, 3, 5, 4]	0.1436
94	[5, 2, 3, 1, 4]	0.1624
81	[5, 4, 1, 2, 3]	0.1786
78	[1, 4, 5, 2, 3]	0.1942
77	[4, 1, 5, 2, 3]	0.2096
75	[2, 3, 1, 4, 5]	0.2246
74	[4, 2, 5, 3, 1]	0.2394
73	[3, 2, 1, 5, 4]	0.2540
71	[5, 2, 1, 3, 4]	0.2682
70	[5, 4, 2, 1, 3]	0.2822
67	[2, 5, 3, 1, 4]	0.2956
66	[2, 5, 1, 3, 4]	0.3088
66	[1, 2, 3, 4, 5]	0.3220
65	[5, 4, 2, 3, 1]	0.3350
64	[2, 5, 4, 3, 1]	0.3478
64	[4, 5, 1, 3, 2]	0.3606
64	[2, 3, 5, 1, 4]	0.3734
63	[4, 5, 3, 2, 1]	0.3860
63	[5, 2, 3, 4, 1]	0.3986
62	[2, 1, 3, 4, 5]	0.4110
60	[5, 1, 2, 3, 4]	0.4230
59	[4, 1, 5, 3, 2]	0.4348
56	[1, 3, 2, 5, 4]	0.4460
56	[1, 2, 4, 5, 3]	0.4572
55	[4, 2, 5, 1, 3]	0.4682
54	[2, 1, 5, 4, 3]	0.4790
54	[4, 1, 2, 5, 3]	0.4898
54	[3, 1, 2, 5, 4]	0.5006

MIP for the Assortment Problem

$$
\begin{aligned}
\max_{\boldsymbol{x},\boldsymbol{W}} \quad & \sum_{i=1}^{m} r_i \cdot \sum_{s=1}^{m} W(i,s,m) & \\
\text{subject to} \quad & W(i,s,t) = (1-\gamma_{t,s}) \cdot W(i,s,t-1) + y_{i,s,t} & \forall i,s,t \\
& y_{i,s,t} \le \gamma_{t,s-1} \cdot W(i,s-1,t-1) & \forall j,s,t \\
& W(t,s,t) = z_{s,t} & \forall s,t \\
& z_{s,t} \le p_{t,s} \cdot \left(\sum_{i=1}^{t-1} \sum_{\ell=s}^{m} W(i,\ell,t-1) + q_{t-1} \right) & \forall s,t \\
& 0 \le q_t \le 1 - x_{\sigma(i)} & \forall i \le t \\
& q_t \ge 1 - \sum_{i=1}^{t} x_{\sigma(i)} & \forall t \\
& 0 \le z_{s,t} \le p_{t,s} \cdot x_{\sigma(t)} & \forall s,t \\
& x_i \in \{0,1\} & \forall i \\
& \sum_{i=1}^{m} x_i \le k &
\end{aligned}
$$

- σ: the ground-truth ranking of the Mallows model;
- r_i: the value of item x_i;
- $\sigma(i)$: the index of the i-th item in the ground-truth ranking σ;
- x_i: indicator variable of whether item x_i is included in the assortment $\mathcal{S}$;
- $W(i,s,t)$: the probability of the i-th item being chosen at position s after t steps of repeatedly insertion;
- $p_{t+1,s}$ is the probability of inserting item the $(t+1)$-th item at position s into a partial permutation of the top t items;
- $\gamma_{t,s}$: $\sum_{\ell=1}^{s} p_{t,\ell}$.

MIP for the Welfare Maximization Problem

$$
\begin{aligned}
\max_{\boldsymbol{x},\boldsymbol{W}^{(1)},\ldots,\boldsymbol{W}^{(n)}} \quad & \sum_{i=1}^{n} p_i \cdot \sum_{j=1}^{m} v_j \cdot \sum_{s=1}^{m} W^{(i)}(i,s,m) & \\
\text{subject to} \quad & \texttt{choiceProbConstraint}(i, \boldsymbol{x}, \boldsymbol{W}^{(i)}) & \forall i \in [n]
\end{aligned}
$$

- $\texttt{choiceProbConstraint}(i, \boldsymbol{x}, \boldsymbol{W}^{(i)})$ refers to the constraint of choice probability between $\boldsymbol{x}$ and $\boldsymbol{W}^{(i)}$ under the i-th Mallows distribution.

Figure 9: MIP formulation of the welfare-maximizing strategy.