CAREER: Behavior-Informed Machine Learning: Improving Robust Learning and Decision Support

Chien-Ju Ho, Washington University in St. Louis

Overview

Machine learning (ML) has seamlessly integrated into various facets of humans' everyday lives, drawing largely from human data for its training. As a result, these ML systems often exhibit and reflect human behavioral biases, leading to concerns in a range of applications, such as social media and medical decision-making. While these concerns underscore the urgent need to consider human behavior when developing ML systems, current methodologies primarily view humans as independent, stochastic data sources or assume they are rational decision-makers. This is despite substantial empirical evidence from psychological studies indicating that human behavior frequently deviates from these models. Such discrepancies highlight the existing gap in integrating empirically-grounded insights on human behavior from psychology into the design of ML systems.

This CAREER project proposes the development of a framework for behavior-informed machine learning that examines and incorporates the impacts of human behavior into the design of machine learning systems. Specifically, I will focus on two key aspects of human behavior in the ML lifecycle: (1) The generation of data used for training machine learning models, and (2) human decision-making in tandem with machine assistance. The proposed research aims to develop ML systems that are robust to behavioral training data and to design assistive ML that enhances human decision-making for improved outcomes. In addition to theoretical contributions, through collaboration with domain experts, the research will be adapted for domain-specific applications such as homelessness prevention and pilot augmentation. This approach ensures that the research findings will have practical relevance in domain applications, promoting their widespread adoption and impact.

Intellectual Merit

The CAREER plan introduces behavior-informed ML, which integrates insights from empirical human behavior into ML design. It not only contributes to our theoretical understanding of how human behavior impacts the design of ML but also offers computationally efficient approaches for incorporating human behavior into ML design. By collaborating with psychology researchers, the plan also delivers a comprehensive empirical understanding of human behavior during ML interactions, advancing both the fields of psychology and ML. This research proposal is interdisciplinary in nature, leveraging techniques from ML, algorithmic economics, optimization, and online behavioral social.

Broader Impacts

The success of the proposed activities will broadly impact a wide range of societal domains. Specifically, incorporating human behavior into ML design offers a way to mitigate growing concerns about deploying biased ML, complementing traditional approaches that impose constraints during ML training. This research holds significant potential for practical applications in the industry. Through our partnership with Boeing, we aim to harness ML for practical purposes, such as enhancing pilot decision-making. In the realm of education, the research aligns with the PI's vision of data-driven personalized education. Furthermore, outreach activities are proposed to disseminate the research to students at various levels and to engage underrepresented groups and undergraduate students in conducting the research.

CAREER: Behavior-Informed Machine Learning: Improving Robust Learning and Decision Support

1 Introduction

Machine learning (ML) has seamlessly integrated into various facets of humans' everyday life, largely deriving its training from human data. Consequently, these **ML systems often exhibit and reflect human behavioral biases, leading to a host of concerns**. A notable example is Microsoft's chatbot, Tay, which was designed to learn from conversations with Twitter users. However, it had to be deactivated after just 16 hours due to its unanticipated adoption of offensive language, a direct consequence of failing to account for human behaviors [90, 158]. Similarly, ML models on social media platforms that decides what content to show to users can inadvertently create echo chambers due to the confirmation bias in user behavior [105, 110, 10]. ML models trained on data generated by doctor annotations might suggest unnecessary treatments due to doctors' action bias towards treating diseases, even when the best course of action might be to wait and observe [156, 37, 94]. Autonomous vehicles designed by learning from human driving behavior could adopt dangerous patterns from aggressive or unsafe drivers [166, 63].

While these examples underscore the pressing need to factor in human behavior when developing ML systems, current ML methodologies mostly either view humans as independent, stochastic data sources [28, 157, 118, 165] or assume that humans are *rational* decision-makers [152, 21, 20, 53, 73, 7], despite the substantial evidence from psychological studies indicating that human behavior frequently deviates from these models [146, 67, 148, 66, 69, 68]. Such discrepancies highlight the existing gap in incorporating empirically-grounded human behavior insights from psychology into the design of ML systems. Furthermore, as the capacity of ML and our understanding of human behavior continue to grow, it also opens up the rich potential of designing ML systems to augment human decision making, especially in high-stakes or ethically-sensitive domains where humans are still desired to be the final decision makers.

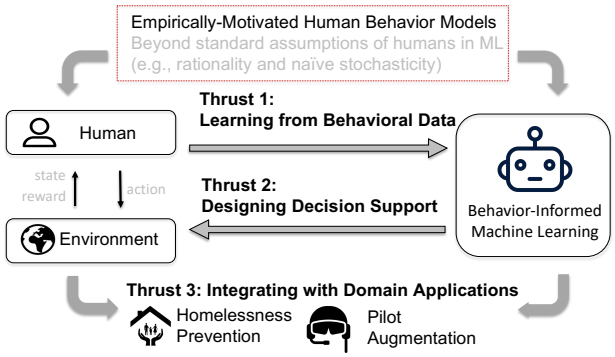


Figure 1: CAREER research plan.

This CAREER project proposes the development of a framework for *behavior-informed machine learning*, which examines and incorporates the impacts of human behavior into the design of ML systems. Specifically, I focus on two key aspects of human behavior in the ML lifecycle: (1) The generation of data used for training machine learning models, and (2) human decision-making in tandem with machine assistance. The proposed research aims to **develop ML systems that are robust to behavioral training data and create assistive ML to augment human decision making for improved outcomes**. In addition to theoretical contributions, through collaborating with domain experts, the research will be adapted to domain applications to ensure the practical relevance and research impact. In more detail, I will investigate the following research thrusts:

- **Thrust 1: Developing foundations for learning from behavioral data.**
  This thrust aims to develop theoretically-grounded foundations for learning from behavioral data, taking into account human behavior during the learning process. Moreover, as ML and generative AI become increasingly integrated into our daily lives, existing behavior models in the literature might not adequately capture human behavior during interactions with ML, and generative AI could also alter the way data is generated from the outset. Therefore, I will conduct behavioral experiments to gain a comprehensive understanding of human behavior in the age of ML. Additionally, I will explore and incorporate the role of generative AI into our framework for learning from behavioral data.

- **Thrust 2: Designing assistive ML to improve human decision making.**
  As ML capabilities continue to advance, there is an increasing need to understand how ML can assist humans in making better decisions. This is especially relevant in high-stakes or ethically-sensitive domains where humans are often preferred as the final decision-makers. In this thrust, I aim to develop assistive ML frameworks that enhance human decision-making by accounting for human behavior. I will delve into when and what kind of support ML should provide, using algorithmic, data-driven, and learning-based approaches. Furthermore, I will conduct behavioral experiments to measure human reliance on ML assistance and design information and explanations to augment the effectiveness of ML assistance.
- **Thrust 3: Integrating with domain applications.**
  While the primary focus of this CAREER plan is on developing a general framework for behavior-informed ML, I will also engage with domain experts to address practical challenges in implementing this framework in specific applications. Specifically, the proposed research will be tailored for use in the areas of homelessness prevention (in collaboration with Prof. Patrick Fowler) and flight pilot augmentation (in partnership with Boeing). This approach ensures that our research findings are not only robust but also practically relevant, fostering their widespread adoption and enhancing their potential impact.

Long-term Goal. My career goal is to develop the foundations for humans and ML to collaborate together and solve problems neither can solve alone. This requires the advancements of ML, the understanding of humans, and the utilization of their interactions. This research proposal serves as a stepping stone to achieving this goal by designing learning algorithms that are robust to human behavior during data generation, and investigating the design of assistive ML to augment humans in making better decisions.

PI Qualifications. The PI has extensive research experience in studying the interactions between humans and ML, using techniques drawn from ML, algorithmic economics, optimization, and online behavioral social science. From the perspective of learning from humans, the PI has explored the problem of eliciting and learning from noisy human-generated data [54, 57, 2, 60, 55, 140, 138, 34, 35] and designing incentives to encourage high-quality data [56, 58, 61, 85]. From the perspective of designing ML to assist humans, the PI's recent works started the exploration of designing ML assistance for specific human models with information and environment design [161, 139, 41, 32] and investigated ethical considerations in leveraging ML in decision making [141, 142, 96, 97]. The PI has extensive experiences in incorporating human behavioral models in ML [138, 161, 143, 142, 41] and understand human behavior in computational environments with behavioral experiments [59, 140, 34, 35, 80, 96, 97]. The PI is active in the research communities. The PI served as the Doctoral Consortium Co-Chair and Works-in-Progress Co-Chair of HCOMP (in 2022 and 2019, respectively), the premier conference in human computation. The PI has also organized workshops at NeurIPS and HCOMP to explore the interactions between humans and machine learning, and served as the area chair, senior program committee, and program committee in major AI/ML conferences.

1.1 Intellectual Merit

This research is interdisciplinary in nature, leveraging techniques from ML, algorithmic economics, optimization, and online behavioral social science. It offers the following intellectual contributions.

Theoretical and algorithmic foundations of behavior-informed ML. This proposal develops the theoretical and algorithmic foundation that integrates empirical insights about human behavior into ML design. While there is broad acknowledgment that human behavior deviates from conventional assumptions, research efforts geared towards incorporating such behavioral insights into ML are still limited. This research addresses the gap in understanding and integrating human behavior in ML design. It will result in theoretically provable and computationally efficient approaches to designing ML systems that learn from human behavior and assist human decisions. The research also offers a novel pathway to mitigate potential concerns related to biased ML deployment by accounting for human behavior, one of the primary sources of ML biases, complementing traditional approaches that impose constraints during ML training.

Empirical understanding of human behavior when interacting with ML. ML is increasingly integrated into human decision-making processes. Recent empirical evidence suggests a noticeable shift in human behavior when ML plays a role in decisions. However, our understanding of human behavior during interactions with ML remains scarce. In collaboration with psychology researchers, this study aims to enrich both the fields of psychology and machine learning by providing comprehensive empirical insights into human behavior *during interactions with ML*. Furthermore, the proposal introduces a behavioral methodology to consider the role of generative AI within the learning framework.

2 Broader Impacts

Societal impacts. The proposed activities will broadly impact numerous societal domains. Specifically, factoring in human behavior in ML design offers a novel and natural method to address concerns regarding the training and deployment of potentially biased ML systems, especially in sensitive areas such as homelessness prevention and medical care. As algorithmic decision-making grows increasingly prevalent in policy-making, this research paves the way for augmenting human decision-making in tackling various societal issues. For instance, the PI has ongoing collaborations with Prof. Patrick Fowler at Brown School of Social Work on homelessness prevention [33], and with Dr. Jason Wellen at the Medical School on applying computational approaches for living donor kidney transplantation [81]. My goal is to continue and grow these collaborations through interdisciplinary efforts at WashU. This will be realized by actively engaging with the Division of Computational and Data Sciences (DCDS), the Center for Collaborative Human-AI Learning and Operation (HALO), and the Transdisciplinary Institute in Applied Data Sciences (TRIADS).

Industry impacts. This research possesses significant potential for practical applications in the industry. I have already collaborated with Boeing to devise strategies that leverage ML to augment pilot decision-making. I have been working with the Office of Industrial Relations at WashU to establish partnerships with other industry players. This includes receiving a faculty research award from JPMorgan and engaging in active discussions with Waters Corporation to implement our research findings in practical settings.

Education and outreach. The CAREER plan includes integrating our research outcomes into realizing my long-term education vision of providing data-driven personalized education. Furthermore, through a range of outreach activities, we aim to disseminate the research outcomes to students at the graduate, university, and high-school levels. The CAREER plan also outlines strategies to actively engage female students and those from traditionally underrepresented backgrounds. I firmly believe – and this belief is integral to my research agenda – that enhancing diversity is fundamental in addressing biases from the outset.

3 Background

This proposal aims to integrate human behavior into the design of ML systems. While the proposed methodologies apply more broadly, the primary focus is in the context of sequential decision making in this proposal. To lay the groundwork for this discussion, we begin with a concise overview of a classical decision-making framework, associated ML frameworks, and existing literature on integrating humans in ML. We then summarize widely recognized human behavioral models from behavioral economics and psychology.

3.1 Decision Making Framework in Machine Learning

We first review Markov decision process (MDP), the classical sequential decision-making framework that serves as the foundation of the proposed research.

Markov decision process (MDP). Markov decision process (MDP) is one of the most standard frameworks for modeling the sequential decision-making environment in ML. An MDP can be characterized by the tuple $\langle S, A, T, R \rangle$, where state space S characterizes the environment a sequential decision maker is interacting with, action space A: actions the decision maker can chose from at each step, state transition function

$T(s'|s,a)$ characterizes how decision maker's actions change the environment, and reward function $R_a(s,s')$ describes the benefits of taking each action.

Reinforcement learning (RL). The standard approach to solve the above MDP and obtain an optimal policy is through reinforcement learning (RL) [65, 137, 92, 93]. The RL agent interacts with an unknown environment and attempts to maximize the total of its collected reward. At each time t, the agent in state $s_t \in \mathcal{S}$ takes an action $a_t \in A$, which returns a reward $R_{a_t}(s_t, s_{t+1})$, and leads to the next state $s_{t+1} \in S$ according to $T(s'|s,a)$, the probability to state s' from s after taking action a. The goal of RL is to learn a policy $\pi(a|s)$ that maximizes the total time-discounted rewards $\mathbb{E}_\pi[\sum_t \gamma^t R_{a_t}(s_t, s_{t+1})|\pi]$, where $\gamma \in (0,1]$ is a discount factor ($\gamma = 1$ indicates an undiscounted MDP). RL has a long history of development, from the seminal Q-learning [154], to more recent deep learning aided approaches [83, 92, 93].

Inverse reinforcement learning (IRL). Inverse reinforcement learning tackles the problem of inferring the reward R from observing the sequence of (s_t, a_t), assuming the observed actions are taken by *rational* decision makers. This problem has also been referred to as apprenticeship learning, or learning by watching, imitation learning etc. Ng et al. [99] is among the first to formalize this problem. The high-level idea is to find a feasible function $R(\cdot)$ such that a_t is the action that maximizes the utility at s_t for all (s_t, a_t) pairs, and impose smoothness constraints on each step's predicted policy to formulate a linear programming problem. Follow-up works [1, 167, 114] have focused on variants of the optimization formulation.

Existing approaches in incorporating humans in ML. Learning from human demonstrations has been extensively studied in weakly supervised learning [15, 98, 125], crowdsourcing [118, 27, 157, 28, 165], and IRL [99, 167, 114]. However, these studies largely assume humans to be rational or naively stochastic, presenting concerns in potentially biased learning outcomes. While some recent research efforts have been devoted to modeling human behavior and integrating it into learning [40, 164, 89, 120, 127], these efforts are limited, and the theoretical understanding of learning from behavioral data remains underdeveloped. The complexity increases as ML and generative AI become more common, affecting both human behavior and data generation. New studies are imperative to comprehend these impacts. The proposed research seeks to bridge this gap by offering a theoretically-grounded foundation, enhancing our grasp of human behavior during interactions with ML, and exploring the influence of generative AI in learning from behavioral data.

For designing ML to assist humans, there has been increasing attention focusing on improving the overall performance of human-ML partnerships [49, 50, 78, 12, 79] and on investigating the interpretability [121, 88, 52, 116, 71, 109] and trustworthiness [29, 36, 30, 86, 159, 160] of ML. However, existing studies have mostly focused on relatively simple one-shot decision making scenarios and often do not incorporate empirically-grounded human behavioral insights in ML design. The goal of this proposal is to explicitly incorporate human behavior in designing assistive ML in sequential decision-making.

3.2 Empirically Motivated Human Behavioral Models

In this CAREER plan, we aim to incorporate empirically grounded human models into the design of ML. While the proposed research is applicable to general human models (including data-driven forms), to make the discussion more concrete, we summarize and provide formulations of some notable classes of human behavioral models in the literature of economics and psychology.

Biased reward evaluation. While it is commonly assumed that humans are rational, taking actions to maximize their expected utility (the expected utility theory [152]), humans are consistently observed to deviate from this assumption. For example, humans often over-estimate small probabilities and react more strongly to losses than gains. The most important theory that summarizes these systematic biases is the Nobel-winning *prospect theory* [66]. Another commonly used theory, also Nobel-winning, is the discrete choice model [91, 131, 144], accounting for the inherent randomness of human decision making by incorporating noises in the utility. These deviations from standard rational assumption can often be captured with humans' biased reward evaluations. Formally, let $(p_1, x_1, ..., p_K, x_K)$ be the *prospect* of an action, where p_k

represents the probability of the outcome x_k happens after taking the action. Let $v(x_k)$ represent the utility of the outcome x_k. The above theories can be summarized as follows.

- Expected utility theory: It predicts that humans will take the action that maximizes $\sum_{k=1}^{K} p_k v(x_k)$.
- Prospect theory: It predicts that humans will take the action that maximizes $\sum_{k=1}^{K} \pi(p_k)u(v(x_k))$, where $\pi(\cdot)$ and $u(\cdot)$ models the humans' distorted interpretations on the probability and utility measure.
- Discrete choice model: It predicts that humans will take the action that maximizes $\sum_{k=1}^{K} p_k v(x_k) + \epsilon$, where ϵ is the additional noise term that incorporates the intrinsic randomness of human decision making.

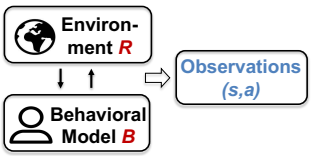
Selected information processing. Humans tend to prioritize certain types of information while neglecting others, leading to decision biases. For example, confirmation bias [102, 72] describes humans' tendency to prioritize information agreeing with their existing beliefs. People exhibiting herding bias [11, 17] prioritize information that aligns with the majority. Anchoring biases [147, 45] describe humans' tendency in prioritizing the initial piece of information they receive. These biases can be formulated by addressing how humans form their subjective beliefs. Let e be the event of interest and I be the provided information. Traditional models assume humans are Bayesian, forming their posterior following the Bayes rule $P(e|I) = P(I|e)P(e)/P(I)$. One way for formulate these biases [117] is through modeling humans' belief updating as $P(e|I) \propto P(I|e)^{\alpha}P(e)^{\beta}P(I)^{\gamma}$ and choosing the values of (α, β, γ) to reflect humans' different weights in different pieces of information during decision making.

Time-inconsistent planning. Humans often cannot reason about future rewards in a consistent manner. For example, humans might be myopic or boundedly rational, failing to properly reason about future rewards due to cognitive and information limitations. Humans might also inherit time-inconsistent reasoning behavior. For example, when choosing between earning \$10 in 100 days or \$11 dollars in 101 days, most people will choose the latter. However, when choosing between earning \$10 now or \$11 tomorrow, many people will change their decisions to the former. This example illustrates *present bias* [111], an common example of humans' time-inconsistent panning. These time-inconsistent biases can be modeled by introducing a *discounting function* $d(t)$ that captures humans' behavior in weighing future rewards. Let R_t denote the expected reward at time t, human's perceptions of the long-term rewards can be modeled as $\sum_t d(t)R_t$. For example, the standard model often set $d(t) = \gamma^t$ with $\gamma \in (0, 1]$. To model myopic or bounded rationality, we can set $d(t) = \gamma^t$ for all $0 \leq t \leq \tau$, and $d(t) = 0$ for all $t > \tau$. For present bias, one common model is hyperbolic discounting: $d(t) = \frac{1}{1+kt}$ for some pre-specified $k > 0$.

4 Proposed Research

4.1 Thrust 1: Developing Foundations for Learning from Behavioral Data

This thrust aims to develop foundations to learn from behavioral data. In this learning setup, we have access to the observations $\{(s_t, a_t)\}$, indicating human action a_t at state s_t, inherently arising from the *interactions* between humans $\mathcal{B}$ and the environment R. For convenience of discussion, we use $\mathcal{B}$ to represent the general human models and abuse the notation R to represent general environment parameters beyond just rewards.[1]



This learning setup poses new challenges compared with traditional ones. First, given the interactive nature, behavioral data often breaks the i.i.d. data assumption made in the vast majority of the ML literature [23]. While some research efforts have attempted to explicitly model human behavior and incorporate the model during learning [40, 164, 89, 120, 127], as shown in my work [138], there exist scenarios in which it is infeasible to learn with any models even with an infinite amount of data. While this challenge has also been reported empirically [123], the theoretical understanding of when learning from behavioral data

[1] The discussion in this thrust also applies to (a simpler setting of) supervised learning, where we observe the feature-label pairs generated by humans $\{(x_n, y_n)\}_{n=1}^{N}$ and aim to uncover the latent mapping from features to labels.

is feasible is still largely lacking. Moreover, as ML and generative AI are increasingly pervasive, existing human models might not be sufficient to capture human behavior with the presence of ML, and generative AI might be involved in the generation of behavioral data.

To address the above challenges, I propose to develop theoretically-grounded foundations of learning from behavioral data (**Task 1.1**), conduct human-subject experiments to understand how human behavior changes when ML is integrated into decision-making (**Task 1.2**), and develop approaches to account for the role of generative AI in the learning framework (**Task 1.3**)

Prior work. The proposed activities in this research thrust will be built on my extensive prior work in crowdsourcing [54, 57, 58, 59, 60, 85, 140, 34, 35], where one key research theme is to infer ground truths from noisy human data. I will extend the standard models assuming humans exhibit zero-mean noises to general human behavioral models. My recent works on incorporating behavioral models motivated by psychology literature in learning frameworks [138, 161, 41] and the experience in conducting human-subject experiments [59, 140, 34, 35, 97] will be the building blocks of the proposed research.

4.1.1 Task 1.1: Developing theoretically-grounded foundations of learning from behavioral data

This task aims to theoretically characterize conditions, in terms of human behavior and environments, for the feasibility of learning from behavioral data. When learning is feasible, I will develop efficient learning algorithms. When learning is infeasible, I will investigate approaches to enable the feasibility of learning by intervening the data generation process.

Theoretically characterizing conditions for feasible learning. For the theoretical characterization, I propose to employ techniques from stochastic approximation [122, 44], modeling the state realizations over time as a random variable, resulting from the interactions between human behavior and the environment. As an illustrative intuition: If a bijection exists between $(\mathcal{B}, R)$ and the state at convergence, we can infer $(\mathcal{B}, R)$ from the converged state, indicating the feasibility of learning (a weaker condition might be sufficient to characterize the learning feasibility). Therefore, by analyzing the convergence and convergence rate of the state trajectory, represented using human behavior and environments, we can identify the conditions of human behavior and environments for the learning feasibility and complexity of learning. After the characterization, I will also examine the *robustness* of the theoretical results. In particular, since no models can perfectly capture human behavior, I will quantify how the deviations in human models propagate to errors in learning through sensitivity analysis and robust design [143].

Designing computationally efficient learning algorithms. Even when learning is feasible, we still need to confront computational challenges due to a larger learning space. To tackle this, I plan to adopt, examine, and compare several approaches. First, I will presume partial access to true environment parameters R_t (e.g., some rewards are known in MDP), e.g., through domain knowledge or historical data, and then implement a two-stage learning process (i.e., infer $\mathcal{B}$ using partial R_t then infer full R with the inferred $\mathcal{B}$). Second, I will impose suitable constraints to reduce the learning space, for instance, by leveraging smoothness constraints or applying domain knowledge on belief models and reward bias models. Lastly, I will resort to sampling and variational techniques, such as Gibbs sampling, to approximate the inference problem.

Taking interventions or increasing diversity during data collection to enable improved learning. One main reason leading to infeasible learning is the human tendency to engage in *exploitation*, resulting in under-represented datasets and feedback loops. Based on this observation, I plan to explore methods to increase the amount of *exploration* in behavioral data during data generation. First, I will quantify the amount of exploration necessary to make learning feasible through an *epsilon-first* approach [145]. I will then aim to reduce the amount of exploration required by strategically deciding when to explore. Second, inspired by recent literature [113, 14] showing that inherent *diversity* could render exploration unnecessary, I will examine the relationship between data diversity (e.g., in terms of population behavior) and the learning efficiency and investigate to what extent increasing data diversity could facilitate more efficient learning.

4.1.2 Task 1.2: Conducting experiments to understand human behavior in the age of ML

In the previous task, I start the investigation by leveraging existing behavioral models from the literature of psychology. Although these models are supported by extensive empirical evidence, they are primarily developed before ML becomes pervasive. Meanwhile, when people become aware that they are interacting with ML, their behavior might shift and existing models might not be sufficient to capture human behavior in the age of ML. For example, my recent work [80] has demonstrated that when humans know their behavior will be used to train ML, compared with not knowing ML is involved, they are more willing to forgo some rewards to ensure the trained ML is fair. As people are getting more aware of the role of ML in our decision making, it is crucial to examine and understand the shifts in human behavior in the age of ML.

Understanding human behavior with the presence of ML. I will conduct experiments to examine how the presence of ML changes human behavior. The results will not only deepen our understanding of human behavior with the presence of ML but also serves as an improved foundation for learning from behavioral data. To conduct the research, following the standard literature, I will start by utilizing social games, such as the ultimatum game [103], dictator game [39], prisoner's dilemma [9], to examine human behavior with the presence of ML. These social games provide succinct abstractions of human behavior and interactions in different contexts and are useful as the starting point towards a comprehensive understanding of humans. I will recruit participants from crowdsourcing platforms, vary the following independent variables in the experiments, and measure human responses as the dependent variables. Standard statistical tests (such as ANOVA and post-hoc t-tests) will be conducted to examine the result significance.

- Whether humans are explicitly interacting with ML. We hypothesize that humans are more likely to care more about ethics (e.g., being fair) when their partners in the game are other humans than ML.
- Whether human decisions will be used to train ML used to play with future players. We hypothesize humans are willing to sacrifice rewards to make the future ML behave in a more *ethical* manner.
- The context of the environment and ML. For example, whether the trained ML will be playing with people they view favorably in the future. Whether the ML training mechanism is known to people.

Collaborating with psychology researchers. For the proposed activities in this task, I will collaborate with Dr. Wouter Kool in the department of Psychology and Brain Sciences at WashU. Dr. Kool and I are currently co-advising a PhD student, Lauren Treiman, with whom we have generated the preliminary result [80] for this task. The proposed research will enable us to obtain a comprehensive understanding of human behavior when ML is integrated into various aspects of decision-making. The outcomes will contribute not only to the ML community but also to the psychology community.

4.1.3 Task 1.3: Incorporating generative AI in learning from behavioral data

The rise of generative AI, such as large language models (LLMs), is significantly reshaping the manner in which we approach problems. In the context of learning from behavioral data, recent research has reported that LLMs outperform human workers in data annotations for certain tasks [47, 38]. Meanwhile, LLMs have also been shown to exhibit biases [3, 112, 42, 26]. Given that we anticipate generative AI will play a pivotal role in shaping behavioral data generation in the near future, in this task, I aim to develop empirical understandings of the *behavior* of generative AI with the aim of accounting for it in our learning framework.

One of the main challenges in understanding generative AI is its black-box nature. In relation to the growing efforts to address the transparency of generative AI [82, 108], I propose adopting a behavioral approach – treating generative AI as a behavioral agent – to understand and incorporate its behavior into our learning framework. It is worth noting that given the rapid and ongoing evolution of generative AI, this task represents a long-term and continuously evolving research agenda. I believe that addressing this is of paramount importance, especially as we approach an era where generative AI might be ubiquitous.

Empirically examine and model the behavior of generative AI. I will begin by characterizing AI behavior using standard human behavioral models, allowing for a direct comparison with human behavior. Our pre-

liminary results suggest that ChatGPT demonstrates certain human-like behavioral patterns, e.g., it reports willingness to sacrifice reward to promote fairness, similar to human choices in our human studies [80]. Considering generative AI is trained on human data, it is useful to understand to what extent it reflects existing human behavioral patterns in different contexts. Moreover, as generative AI might exhibit distinct behavioral characteristics, I will develop new models with explainable insights for AI's decision-making processes leveraging recent efforts in behavioral modeling with computational approaches [18, 107].

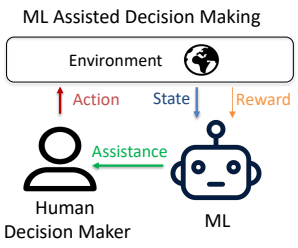
Theoretically characterize the capacity of generative AI. Building on the understanding of the behavior of generative AI developed above, I will utilize the theoretical results from Task 1.1 to assess whether learning is feasible using data from generative AI *alone* in the given task domain. If feasible, it implies that generative AI has sufficiently encoded the task domain in its model, potentially suggesting that human involvements might not be needed in solving tasks in that specific domain. In addition to the traditional measure of empirical accuracy (e.g., the empirical ratio of data aligning with the partial gold-standard dataset), I expect the results would correlate to our diversity discussion in Task 1.1, e.g., if there exists diverse data sources in the task problem domain and if the hyper-parameters (e.g., *temperature*) of generative AI leads to more diverse output, it might be more likely to enable learning with generative AI alone. This approach provides a potential venue for us to characterize the capacity of generative AI.

Leveraging both humans and generative AI to achieve efficient learning. In instances where learning is infeasible using solely AI-generated data, I will examine methods to incorporate humans to enable learning. I intend to also draw from approaches developed in Task 1.1. Specifically, by viewing generative AI as a behavioral agent with distinct behavioral patterns, I will investigate which types of supplemental behavioral data from humans are most likely to enable effective learning. This will be guided by our findings on the relationship between *diversity* in the dataset and the efficiency and feasibility of learning.

4.1.4 Expected outcomes

The successful completion of Thrust 1 will establish a theoretically-grounded foundation for learning from behavioral data. Importantly, this learning foundation will factor in the increasing prevalence of ML and generative AI in human decision-making and the generation of behavioral data. It offers a framework to combine data from both humans and generative AI, paving the way for more efficient learning. Empirically, we also anticipate gaining profound insights into human behavior with the presence of ML – a relatively under-explored area – which contributes to both the fields of psychology and ML.

4.2 Thrust 2: Designing Behavior-Aware Assistive ML to Improve Human Decision-Making



Humans often make suboptimal decisions and engage in "on-the-job-training," i.e., learn to make better decisions while making these decisions [115, 132, 13]. Conversely, the rapid advancements in ML highlight its potential to enhance human performance and expedite their learning with ML assistance. In this research thrust, our aim is to develop a framework for ML-assisted decision-making that accounts for human behavior. The goal is to determine when and what kind of assistance ML should offer, especially when only a limited number of interventions are feasible. In this framework, ML provides guidance to humans, who subsequently make the final decisions. This setup is particularly relevant in high-stakes or ethically-sensitive domains where humans are favored as the final decision-makers. This framework also holds promise in education, pinpointing the most impactful state-action pairs to improve decision outcomes (e.g., Section 5.1).

While there have been recent research efforts to leverage ML to enhance human decisions [49, 50, 78, 12, 160, 87], most studies have focused on simpler one-shot settings and often do not incorporate empirical behavioral insights in the ML design. Designing assistance *policies* that addresses sequential decisions while taking into account human behavior remains largely under-explored. This thrust aims to explore the design of ML assistance policies tailored for sequential decision-making, factoring in human behavior. Apart from

human decision-making behavior (mostly referred to as human behavior in our discussion), another crucial aspect of human behavior in this framework is humans' reliance on ML assistance. The factors influencing a human's choice to follow ML assistance are varied and is not well understood. In this thrust, I will start by assuming knowledge of human reliance and develop a framework to design ML assistance policies that factor in human decision-making behavior (**Task 2.1**). Subsequently, I will seek a comprehensive empirical understanding of human reliance in sequential decision-making through behavioral experiments (**Task 2.2**). Finally, I will leverage this empirical knowledge to study the design of explanations accompanying ML assistance, aiming to enhance human understanding and reliance on such assistance (**Task 2.3**).

Preliminary: The ML assistance framework. To make the discussion more concrete, suppose the human decision maker is solving a decision making problem formulated as an MDP. Let the human decision-making policy be $\pi_{\mathcal{B}}(a|s)$, representing the probability for the human to choose action a at state s without assistance. The goal of ML is to maximize the total rewards derived from human actions by providing assistance. ML assistance can be *pushed*-based, where ML decides when to assist and push it to humans, or *pull*-based, where ML passively responds to human-initiated requests. We can formulate the differences with $P_{req}(s)$, representing the probability for humans to ask for assistance. For pushed-based assistance, $P_{req}(s) = 1 \ \ \forall s$, and ML decides whether to provide assistance. We will address both cases with known and unknown $P_{req}(s)$. Let ML's assistance policy be $\rho(a|s)$, the probability for ML to suggest action a at state s, and $\theta(s)$ be the *reliance policy* denoting whether the human adopt ML assistance at state s. Now let $(\pi_{\mathcal{B}} \oplus \rho \oplus \theta \oplus P_{req})(a|s)$ be the final human policy with ML assistance, determined by $\pi_{\mathcal{B}}, \rho, \theta$,$P_{req}$ jointly. For example, for pushed-based ML assistance with $P_{req}(s) = 1 \forall s$, we might have $(\pi_{\mathcal{B}} \oplus \rho \oplus \theta \oplus P_{req})(a|s) = (1-\theta(s))\pi_{\mathcal{B}}(a|s) + \theta(s)\rho(a|s)$. The ML's assistance design problem is then to choose the assistance policy $\rho(a|s)$ to maximize the total expected reward subject to the pre-defined constraints of ML assistance policy. One natural example of the constraint would be to ensure ML does not intervene with human decision-makers too much, i.e., the distance $D(\pi_{\mathcal{B}}, \rho)$ between human policy and ML policy is small for some distance measure D (such as ℓ_p-norms).

Prior work. The proposed activities in this thrust are grounded in my extensive prior work. The problem formulation aligns with a Stackelberg game formulation, where ML determines the policy for providing assistance, and humans decide their course of action based on this assistance. I have explored the optimization problems associated with Stackelberg games across various domains [61, 143, 32, 142, 41, 161]. My expertise in bandit learning [61, 85, 138, 142] and my recent investigations into the role of ML in assisting human decision-making [96, 97] also serve as the technical foundations for this thrust.

4.2.1 Task 2.1: Developing efficient approaches for designing ML assistance

This task aims to design ML assistance policies that account for human behavior (in decision making), assuming human reliance is known. I will address situations of varying environment complexities and informational assumptions by leveraging algorithmic, data-driven, and learning approaches. The expected outcomes of this task are characterizations of the complexity involved in designing ML assistance under different human behaviors and environments, with the associated approaches for identifying the policy.

Solving the bi-level optimization for low-complexity environments and known human behavior. I will start with known human behavior and low-complexity environments, where the *optimal* policy for the environment can be algorithmically derived, e.g., using value iteration. Even in this simplest setting, my prior work [161] has proved that this problem is a *bi-level optimization* problem, since the design of assistance needs to account for its impacts to human decisions, and is NP-hard to solve in general. To address this, I consider two cases. When the human models are differentiable (e.g., discrete choice model), I will leverage first-order methods and characterize the computational complexity and convergence to the optimal solution with different human models. When the human models are not differentiable (e.g, expected utility theory), I propose to utilize techniques from my work in algorithmic information design [32, 41] that

utilize the duality theory to characterize the properties of the optimal solution, which help identify conditions for computationally feasible solutions to exist, and derive the corresponding algorithms when it is feasible.

Data-driven approaches for complex environments For complex environments, I propose to leverage data-driven approaches motivated by self-play [129, 31, 162]. The main difference is that we need to account for human behavior after ML assistance. Specifically, I will utilize a neural network to represent the assistance policy. For each assistance policy, by applying the given human model, we can evaluate its performance through simulations. To optimize the assistance policy, I can then apply gradient descent: Draw problem instances from a pre-specified distribution and perform stochastic gradient descent to maximize the policy performance (applying soft-max for non-differentiable objectives). I will examine the empirical performance of this approach with different settings of human models, environments, and data distributions.

Bandit learning approaches for unknown human behavior. The above approach requires the knowledge of human behavior to simulate policy performance in each iteration. To address the setting when human behavior is unknown, I will take an online learning approach, designing ML to interact with human decision makers in the environment and adaptively update the assistance policy based on observations. This leads to the the classical exploration-exploitation trade-off, commonly addressed in the bandit problem [77, 8, 22]. The main challenge of applying bandit in this setting is that the space of arms (i.e., the space of assistive policies) is large/infinite and could require too many explorations for bandit algorithms to be useful. I plan to leverage domain knowledge of the problem structure (abstracting key properties of human behavior and environments) to parameterize the arm space and utilize the technique from my prior work [61] to deal infinite but parameterized arms. The goal is to develop bandit algorithms for the problem and characterize the theoretical guarantees for different models of human behavior and environments.

4.2.2 Task 2.2: Understanding human reliance on ML assistance with human-subject experiments

The previous task has assumed that human reliance on ML assistance, $\theta(s)$, is known and given. However, in practice, appropriately formulating $\theta(s)$ is not trivial and not well understood. While there has been a growing line of literature aiming to understand humans' trust and reliance on ML [160, 87, 119, 163, 84, 151], including my work [97], existing studies mostly focus on the one-shot decision making scenarios, and limited is known about humans' reliance on ML assistance in sequential decision making.

Empirically examining human reliance on ML assistance. In collaboration with Dr. Ming Yin, a leading HCI expert in human trust and reliance, I will conduct randomized human-subject studies to understand how humans' reliance on ML is influenced by various factors under the sequential decision making setting. In particular, consistent with theoretical models previously proposed for human-automation interaction [62, 124], I expect humans' adoption of ML advice under sequential decision making settings can be influenced by factors related to *humans*, *ML*, and the *environment*. I will conduct experiments to understand:

- How factors related to ML, including the presentation format of ML recommendations, the provision of ML explanations, and the human-likeness of ML, influence humans' adoptions of ML assistance?
- How factors related to the decision making environment, including the variability and complexity of the environment, influence humans' adoptions of ML assistance?
- How factors related to humans, including their risk attitudes, their value similarity with ML, and their subjective perceptions of ML trustworthiness, influence humans' adoptions of ML assistance?

General experimental designs. I will start by designing experiments that only a single independent variable varies for each. That is, different experimental treatments will be created corresponding to different "levels" of the independent variable (e.g., timing of ML recommendations). For dependent variables, I will record whether human participants decide to rely on the ML's assistance to estimate $\theta(s)$, as well as their final decision making performance. To align with the AI trust literature, I will also ask human participants to self-report their perceived trust level in ML both at a fixed interval and at the end of the experiment. After collecting the measurements on all the dependent variables, we can conduct statistical tests across treatments

to examine if the independent variable varied in the experiment affects decision makers' adoption of ML assistance and subjective trust on ML, decision making performance, and learning outcome. Moreover, additional experiment can be carried out to vary multiple independent variables simultaneously, which will allow us to understand how they interact with one another to affect the dependent variables of interests.

4.2.3 Task 2.3: Designing information and explanations to accompany ML assistance

In our framework, ML determines when and what kind of assistance to offer to human decision-makers, with the assistance typically in the form of suggested actions. While this captures a general form of the ML-assisted decision-making paradigm, in practical situations, it's often valuable for ML to provide humans with additional information and explanations. For instance, ML might suggest actions that lead to a lower immediate payoff but offer greater long-term rewards (e.g., recommending regular exercise for long-term health benefits). Without an understanding of the rationale behind ML's assistance, humans might be skeptical of the suggestions, making the assistance less effective.

Designing information to accompany ML assistance. If humans consistently ignore ML assistance, the ML's capability to aid them becomes limited. In this task, we explore the potential of information design to encourage human decision-makers to follow the provided ML assistance. Drawing from my prior work in information design [139, 41, 32], I will formulate the problem where the assistive ML needs to decide what information to present to human decision makers to encourage them to follow the assistance. To address this problem, we need to model how humans respond to presented information. I will begin with standard models, where humans incorporate provided information to form a posterior, and make decisions that maximize their utility. I will then extend such models to consider general behavioral models leveraging our empirical understanding developed in Task 2.2. I will design algorithms for finding the information to present with respect to the different models of human behavior. Moreover, since provided information would influence humans' reliance on ML assistance (encoded in $\theta(s)$), we will incorporate these back into the ML assistance framework to explore the joint design of ML assistance and accompanying information.

Designing explanations for the assistance policy. We now explore the design of explanations aimed at enhancing user understanding in the assistance policy. I will employ techniques from the explainable AI planning (XAIP) literature [100, 101, 149]. In XAIP, the typical objective is to explain to a user why a specific plan is feasible or optimal. In our case, the goal shifts to clarifying why the assistance policy possesses certain trustworthy characteristics, drawing on our empirical findings from Task 2.2. To design these explanations, we will use probabilistic logical inference methods [150] to address the main challenge of presenting personalized explanations. I will provide personalized explanations through delivering explanations at various abstraction levels tailored to users with differing expertise and user models (e.g., characterized in Task 2.2). Moreover, I will also combine model reconciliation [25, 24, 134, 135] with contrastive explanation [51, 70, 126] techniques so that explanations that bridge the two models also clarify the foils provided by users. The efficacy of explanations will be assessed by measuring users' self-reported understanding of the assistance policy, perceived trust, their reliance on ML assistance, and overall decision performance.

4.2.4 Expected outcomes

Upon successful completion of this thrust, we will have established a framework for designing ML assistance policies tailored for sequential decision-making, factoring in both human decision-making behavior and human reliance on ML. This framework also provides accompanying information and explanations to improve human understanding of the ML assistance policy. Additionally, we will gain deeper insights into human reliance on ML assistance in sequential decision-making, a currently under-explored research area.

4.3 Thrust 3: Integrating with Domain Applications

In Thrusts 1 and 2, the goal is to develop a framework for behavior-informed ML, incorporating human behavior in the design of ML systems. While the framework is intended to be general, deploying the

framework in domain applications may introduce various domain-specific challenges. For instance, when allocating scarce societal resources for homelessness prevention, it is important not only to maximize the effectiveness of these resources but also to ensure that the allocation of resources is *fair and equitable* across different social groups. When designing decision support systems for airplane pilots, in addition to maintaining decision efficiency, *safety* is of the utmost importance.

In this thrust, I aim to collaborate with domain experts to tackle practical challenges when deploying this framework in ethically-sensitive or high-stakes domains, where humans are still desired to be the final decision-makers. In particular, the proposed research will be tailored for use in an ethically-sensitive domain of homelessness prevention (with Prof. Patrick Fowler at the Brown School of Social Work) and a high-stakes domain of flight pilot augmentation (with Boeing). In the long term, I plan to harness the interdisciplinary efforts at WashU to expand this research into other application domains, including the Division of Computational and Data Sciences (DCDS), the Center for Collaborative Human-AI Learning and Operation (HALO), and the Transdisciplinary Institute in Applied Data Sciences (TRIADS).

4.3.1 Task 3.1: Domain application: Data-driven decision support for homelessness prevention

This task extends my existing collaboration [33] with Prof. Patrick Fowler to develop data-driven decision support for homelessness prevention. The problem of homelessness, a longstanding societal issue, presents significant personal and communal repercussions. Local systems dedicated to addressing homelessness often face a scarcity of resources, making it challenging to fulfill the demand for housing support. The decision-making process is further challenged by the sequential nature, e.g., the resource allocated will not be available for some uncertain period of time. The current decision-making processes for distributing these resources are largely unexplored [19, 43, 128], leaving room for improvement in terms of both efficiency and equity. This opens up two important research directions that align with this proposal: (1) Learn from past data to understand the impacts of resource allocation, thereby allowing us to derive insights to optimize future decisions. (2) Provide decision support for human decision makers in deciding the resource allocation.

Account for human behavior when learning from past data. There is a growing effort to use data-driven approaches to inform decision-making policies in homelessness prevention [46, 74, 76]. Specifically, Prof. Fowler has been involved in the St. Louis Regional Data Alliance [6], an initiative that aims to curate community data to improve community health. Building on this effort, Dr. Fowler and I have been co-advising a PhD student, Alex DiChristofano, in conducting preliminary analyses of St. Louis regional data in homelessness prevention. We have identified two types of human behavior that could inject biases into the data. The first comes from the recipients of resources. In homelessness prevention, when people seek help, they are not immediately assigned resources due to the resource scarcity. Instead, they are placed on a waitlist and only receive resources when resources become available. This waiting process creates unequal *drop-out* rates across social groups, e.g., we found that females are more likely to leave the system before resources become available. Failure to account for this drop-out inequality could lead to biased learning outcomes. The second type of behavior comes from the parties (e.g., social workers) that decide how to allocate resources. While there are general guidelines in the decision-making policy, the past data largely reflects the decision-makers' judgments. In this task, we aim to identify and incorporate these types of human behavior during the training of ML based on past data.

Designing decision support. In the decision of allocating resources for homelessness, there is no clear right or wrong answer. Social workers often need to balance multiple ethical principles, such as prioritizing outcomes (reducing homelessness) or prioritizing the most vulnerable individuals [75]. When designing decision support systems, we must consider decision-makers' preferences and constraints. In this task, we will work with local homelessness service providers, the St. Louis Area Regional Commission on Homelessness (SLARCH) – a nonprofit organization that coordinates homeless service provision across the St. Louis region. By conducting qualitative surveys and interviews, we aim to gain better insights into their

decision-making process, their objectives in decision-making, and the types of decision support needed to inform the design of our assistive ML. Furthermore, we will work with social workers, the decision-makers in the field, recruited through SLARCH, to evaluate and deploy our research.

4.3.2 Task 3.2: Domain application: Decision support for airplane pilots

This task aims to launch our newly initiated collaboration with Boeing in designing decision support for pilot decision-making. In this application domain, in addition to decision efficiency, safety is of paramount importance. To make the discussion more concrete, we will discuss the design of pilot augmentation to address runway incursions – a significant aspect of runway safety. Runway incursion [5] refers to an incident involving an incorrect presence of an aircraft, vehicle, or person on a runway designated for take-off or landing. In severe cases, runway incursions could lead to tragic events. Given the gravity of this problem, there has been research devoted to avoiding such incursions, including accident prediction [136, 130, 48] and system design to detect obstacles and alert pilots [64, 104, 106, 155]. Meanwhile, the Federal Aviation Administration (FAA) have reported that pilot behavior is involved in 65% of all runway incursions [4]. Therefore, in this task, we plan to adopt a behavior-informed approach in addressing the runway incursion problem. We will examine existing datasets and behavioral data from simulated platforms to identify pilot behavioral patterns in the context of runway incursions. Moreover, we will design decision support that provides interventions to prevent runway incursion events.

Proposed research. For the question of learning from behavioral data, we will tap into two data sources. The first is the public ASRS (Aviation Safety Reporting System) dataset, FAA's voluntary confidential reporting system that collects confidential reports of near misses or close call events to enhance aviation safety. Using this public dataset, we can pinpoint generic characteristics of runway incursions. Next, we will use the flight simulator X-Plane, acquired by WashU in a prior collaboration with Boeing, to gather individual behavioral data. This will help in identifying personalized behavioral patterns in runway safety. After identifying the behavioral patterns, we will design decision support systems that aim to maximize decision efficiency (e.g., time for departing/landing) while maintaining safety constraints. The study will initially be conducted in an academic setting, recruiting from the general population (e.g., college students) to operate the flight simulator. After obtaining the preliminary results, in collaboration with Boeing, the valuations will be carried out with domain experts and real pilots using surveys and simulations. The study will also be broadened to other contexts like inflight weather encounters and wake turbulence encounters.

4.4 Evaluation Plan

The proposed research will span five years. The tasks in Thrusts 1 and 2 are organized sequentially. Meanwhile, Thrust 3 will be conducted concurrently with the first two thrusts, leveraging results from them and providing feedback on design challenges. For the evaluations, there are three main components:

- Algorithm and theory: Throughout the proposal, I will develop new algorithms and theories. To evaluate our results, I will derive the performance guarantees (regret bounds or convergence rate) and analyze the computational complexity of the proposed algorithms. Simulations and human-subject experiments will be performed to evaluate the algorithm performance under the conditions both when users follow our proposed models and when users do not exactly follow to test for robustness of our proposed algorithms.
- Data collection: Task 1.2 and 2.2 involve collecting data using human-subject experiments. With collaborations with experts in psychology and HCI, I will follow the best practice in conducting the experiments, including pre-registering the hypothesis and performing appropriate statistical tests (e.g., ANOVA, post-hoc t-tests, mixed effects model). The collected data will be made publicly available to the research community. I believe the large-scale behavioral data would be of important research value.
- Deployment: For tasks in Thrust 3, I aim to deploy the proposed research in domain applications. In addition to the evaluations above, I will work with domain experts to develop our evaluation plan and solicit feedback of the proposed framework through interviews and surveys.

5 Education Plan

5.1 Towards Data-Driven Personalized Education: A Behavior-Informed Approach

My long-term vision in education is to develop data-driven personalized education, i.e., designing personalized curriculum and assistance that improves individual learning with data-driven approaches. This vision aligns with this CAREER plan on designing ML that learns from human behavioral data and assisting humans. As a starting point to realize this vision, I have started to conduct research in the domain of Chess to develop personalized ML assistant. In collaboration with Kassa Korley, who holds the title of International Master and was the youngest African American to earn the title of National Master in the US, we have investigated the question of curriculum design, i.e., what set of moves should be provided to assist Chess players based on their skills, using data-driven approaches. In particular, leveraging the abundant amount of human play data in online Chess platforms (Lichess.org), we have developed ML models that can identify human behavior patterns at different skill levels. We then leveraged both the ideas of designing ML assistance in this proposal and curriculum learning [16, 153, 133] to identify the curriculum most likely to improve players with given skill levels. Our preliminary results [95], showing that the approaches identify curricula that align with domain knowledge and improve win rate, holds potential in designing personalized tool to improve human learning in Chess. We have recently obtained WashU IRB approval to recruit chess players to examine the effectiveness of our approaches in practice.

I will also collaborate with Prof. Dennis Barbour, through co-advising a PhD student, Robert Kasumba, to extend the approach to mathematical education. Prof. Barbour has employed data-driven approaches to explore the connection between students' mathematical learning skills and general executive function skills, such as cognitive flexibility, working memory, and inhibitory/attentional control. The goal of this collaboration is to improve personalized education in the setting of improving students' mathematical skills.

5.2 Course and Teaching Development

The research goal of the PI is to combine the strengths of both humans and machine learning (ML) to solve tasks neither can solve alone. To achieve this goal, we need to advance our understanding of ML, humans, and the interactions between them. Correspondingly, the education goal of the PI is to prepare students on these fronts. The PI plans to introduce a new graduate-level course *Human-AI Interaction and Collaboration*. In addition to the general coverage of ML and human modeling (from behavioral economics, psychology, and HCI), there will be two main themes for the course topics. First. we will cover and discuss human-in-the-loop machine learning, addressing the techniques of incorporating humans in the learning process to advance machine learning. Second, we will discuss topics with a human-centered focus, including how humans process information from ML (such as interpretability, trustworthiness, and topics explores in this proposal) and how ML impacts human welfare (such as fairness, privacy, and ethical concerns). We will also include domain applications in social sciences and healthcare in the course materials (in the form of assignments, projects, or guest lectures) by leveraging the Division of Computational and Data Science (DCDS) and the Center for Collaborative Human-AI Learning and Operation (HALO) at WashU.

5.3 Outreach to High-School Teachers and Students

I will work with the Institute for School Partnership (ISP) at WashU to design outreach activities for high-school teachers and students. The goal is to provide professional developments for teachers and broaden the dissemination of research, and to cultivate next-generation scientists through exposing students to academic research. In particular, we will work with the ISP for the *Teacher-Researcher Partnership*, under which teachers work in the faculty's lab for 4-6 weeks, learning and translating research ideas into lessons at grade level. We plan to host one teacher in each of the first two summers. Based on the partnership outcomes, we will participate in the *Hot Topic Series* at ISP and invite 20 high-school teachers each year from year 3 to 5 to disseminate the research and curriculum design to maximize the potential outreach.

I will host an annual one-day summer workshop "Human-Centered Machine Learning" for local high-school students of low-income background. I will join force with existing efforts at the McKelvey School of Engineering, which in Summer 2022 has conducted a summer camp that is now planned to be an annual event. The workshop will include an overview of ML and human behavior and engage students in group projects guided by undergraduate TAs. We will prepare datasets and ML modules for students to explore the impact of human behavior in the design of ML, both for how human behavior leads to learning (how biased dataset leads to biased learning outcome) and how ML can assist humans in overcoming biases.

5.4 Engaging Underrepresented and Undergraduate Students

I am committed to recruiting female and underrepresented minority (URM) students. Among my 5 PhD students, one is female and one is African American. I have also worked with 6 female and URM undergraduate/master students (out of 13 students I have worked with) at WashU so far. Among the 6 students, four have continued their graduate studies after graduation (at Stanford, Duke, Penn State, and Cornell), one went to the industry (at Google), and one just enters his senior year. To continue this commitment, I will leverage the institutional effort. The CSE department is committed to the goal of increasing the representation of women at the Ph.D. level. e.g., through funding a Platinum Sponsorship of Grace Hopper and hosting various events. I will also work with WashU Summer Engineering Fellowship (WUSEF), which provides funds for students from backgrounds underrepresented in the STEM fields to perform summer research, and the Missouri Louis Stokes Alliance for Minority Participation (MOLSAMP), of which WashU is a participating institution, for offering summer research opportunities for minority participation.

Undergraduate students will be heavily engaged in the proposed research. I am actively involved in the WashU NSF REU site "Big Data Analytics". The 5 students I advised at the REU site have all continued their graduate studies (at UT Austin, Duke, CMU, Yale, and Cornell) after graduation. I am committed to annually support REU/WUSEF research projects during the summer and individual research during academic year inspired by this proposal, e.g., understanding human behavior through experiments or analyzing datasets.

5.5 Evaluation Plan

The educational efforts will be continuously assessed in collaboration with my partners. For data-driven personalized education, I will first employ simulations to assess whether the proposed methods enhance the performance of ML models trained on human data. Subsequently, after securing IRB approval (we have recently been approved for the Chess domain), I will examine the effectiveness of our strategies on human learning. This will be gauged based on participants' performance in the domain (e.g., win rate in Chess).

I have allocated budgets to collaborate with the ISP for evaluating activities aimed at broadening research participation, as well as with the Center for Integrative Research on Cognition, Learning, and Education for assessing the proposed course. The evaluations for the proposed course will be conducted based on multiple metrics, including whether students obtain firm grasp of the subject (by constructing a knowledge inventory) and whether the course motivates students in applying the knowledge in different domains. For broadening research participation, I will conduct anonymous surveys to high-school teachers/students before and after the event to evaluate their understanding of the topic and their aspirations in pursuing higher-education in STEM. For research engagements with URM, female, and undergraduate students, I will conduct two interviews (before and after) to identify potential areas of improvements.

6 Results from Prior NSF Support

Dr. Ho is a co-PI on "FAI: FairGame: An Audit-Driven Game Theoretic Framework for Development and Certification of Fair AI" (IIS-1939677, $444,145, Jan 2020 to Dec 2023). *IM*: This project provides a general framework for fair decision making and auditing in stochastic, dynamic environments. PI Ho has published six publications in this project [96, 143, 34, 142, 35, 97]. *BI*: The work supports the training of graduate students and the development of new auditing algorithms that have impacts to AI and society.

References

[1] Pieter Abbeel and Andrew Y Ng. Apprenticeship learning via inverse reinforcement learning. In *Proceedings of the twenty-first international conference on Machine learning*, page 1. ACM, 2004.

[2] Jacob Abernethy, Yiling Chen, Chien-Ju Ho, and Bo Waggoner. Low-cost learning via active data procurement. In *16th ACM Conf. on Economics and Computation (EC)*, 2015.

[3] Abubakar Abid, Maheen Farooqi, and James Zou. Persistent anti-muslim bias in large language models. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, pages 298–306, 2021.

[4] Federal Aviation Administration. Runway incursion totals for FY 2014. `http://www.faa.gov/airports/runway_safety/statistics/regional/?fy=2014`, 2014.

[5] Federal Aviation Administration. Runway incursions. `https://www.faa.gov/airports/runway_safety/resources/runway_incursions`, 2022.

[6] St. Louis Regional Data Alliance. Community information exchange. `https://stldata.org/project-community-information-exchange/`, 2021.

[7] Tal Alon, Magdalen Dobson, Ariel Procaccia, Inbal Talgam-Cohen, and Jamie Tucker-Foltz. Multiagent evaluation mechanisms. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020.

[8] Peter Auer, Nicolò Cesa-Bianchi, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, 47(2-3):235–256, 2002. Preliminary version in *15th ICML*, 1998.

[9] Robert Axelrod. Effective choice in the prisoner's dilemma. *Journal of conflict resolution*, 24(1): 3–25, 1980.

[10] Eytan Bakshy, Solomon Messing, and Lada A Adamic. Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132, 2015.

[11] Abhijit V Banerjee. A simple model of herd behavior. *The quarterly journal of economics*, 107(3): 797–817, 1992.

[12] Gagan Bansal, Besmira Nushi, Ece Kamar, Walter S Lasecki, Daniel S Weld, and Eric Horvitz. Beyond accuracy: The role of mental models in human-ai team performance. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, 2019.

[13] Hamsa Bastani, Osbert Bastani, and Wichinpong Park Sinchaisri. Improving human decision-making with machine learning. *arXiv preprint arXiv:2108.08454*, 2021.

[14] Hamsa Bastani, Mohsen Bayati, and Khashayar Khosravi. Mostly exploration-free algorithms for contextual bandits. *Management Science*, 67(3):1329–1349, 2021.

[15] Shai Ben-David, Dávid Pál, and Shai Shalev-Shwartz. Agnostic online learning. In *COLT*, volume 3, page 1, 2009.

[16] Yoshua Bengio, JÃ©rÃ´me Louradour, Ronan Collobert, and Jason Weston. Curriculum learning. In *Proceedings of the 26th annual international conference on machine learning*, 2009.

[17] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, 100(5):992–1026, 1992.

[18] David D Bourgin, Joshua C Peterson, Daniel Reichman, Stuart J Russell, and Thomas L Griffiths. Cognitive model priors for predicting human decisions. In *International conference on machine learning*, pages 5133–5141. PMLR, 2019.

[19] Molly Brown, Camilla Cummings, Jennifer Lyons, Andrés Carrión, and Dennis P Watson. Reliability and validity of the vulnerability index-service prioritization decision assistance tool (vi-spdat) in real-world implementation. *Journal of Social Distress and the Homeless*, 27(2):110–117, 2018.

[20] Michael Brückner and Tobias Scheffer. Stackelberg games for adversarial prediction problems. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 547–555, 2011.

[21] Michael Brückner, Christian Kanzow, and Tobias Scheffer. Static prediction games for adversarial learning problems. *Journal of Machine Learning Research*, 13(Sep):2617–2654, 2012.

[22] Sébastien Bubeck and Nicolo Cesa-Bianchi. Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems. *Foundations and Trends in Machine Learning*, 5(1):1–122, 2012.

[23] Longbing Cao. Non-iidness learning in behavioral and social data. *The Computer Journal*, 57(9): 1358–1370, 2014.

[24] Tathagata Chakraborti, Sarath Sreedharan, Yu Zhang, and Subbarao Kambhampati. Plan explanations as model reconciliation: Moving beyond explanation as soliloquy. In *IJCAI*, pages 156–163, 2017.

[25] Tathagata Chakraborti, Sarath Sreedharan, and Subbarao Kambhampati. Balancing explicability and explanations in human-aware planning. In *IJCAI*, pages 1335–1343, 2019.

[26] Yang Chen, Meena Andiappan, Tracy Jenkin, and Anton Ovchinnikov. A manager and an ai walk into a bar: Does chatgpt make biased decisions like we do? *Available at SSRN 4380365*, 2023.

[27] Sharath R. Cholleti, Sally A. Goldman, Avrim Blum, David G. Politte, and Steven Don. Veritas: Combining expert opinions without labeled data. In *Proceedings 20th IEEE international Conference on Tools with Artificial intelligence (ICTAI)*, 2008.

[28] A. P. Dawid and A. M. Skene. Maximum likeihood estimation of observer error-rates using the EM algorithm. *Applied Statistics*, 28:20–28, 1979.

[29] Berkeley J Dietvorst, Joseph P Simmons, and Cade Massey. Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1):114, 2015.

[30] Berkeley J Dietvorst, Joseph P Simmons, and Cade Massey. Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science*, 64(3): 1155–1170, 2018.

[31] Anthony DiGiovanni and Ethan C Zell. Survey of self-play in reinforcement learning. *arXiv preprint arXiv:2107.02850*, 2021.

[32] Bolin Ding, Yiding Feng, Chien-Ju Ho, Wei Tang, and Haifeng Xu. Competitive information design for pandora's box. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 353–381. SIAM, 2023.

[33] Zehao Dong, Sanmay Das, Patrick Fowler, and Chien-Ju Ho. Efficient nonmyopic online allocation of scarce reusable resources. In *Proceedings of the 20th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2021.

[34] Xiaoni Duan, Chien-Ju Ho, and Ming Yin. Does exposure to diverse perspectives mitigate biases in crowdwork? an explorative study. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, pages 155–158, 2020.

[35] Xiaoni Duan, Chien-Ju Ho, and Ming Yin. The influences of task design on crowdsourced judgement: A case study of recidivism risk evaluation. In *Proceedings of the ACM Web Conference 2022*, pages 1685–1696, 2022.

[36] Mary T Dzindolet, Linda G Pierce, Hall P Beck, and Lloyd A Dawe. The perceived utility of human and automated aids in a visual detection task. *Human Factors*, 44(1):79–94, 2002.

[37] Joann G Elmore and Suzanne W Fletcher. Overdiagnosis in breast cancer screening: time to tackle an underappreciated harm. *Annals of internal medicine*, 156(7):536–537, 2012.

[38] Zohar Elyoseph, Dorit Hadar-Shoval, Kfir Asraf, and Maya Lvovsky. Chatgpt outperforms humans in emotional awareness evaluations. *Frontiers in Psychology*, 14:1199058, 2023.

[39] Christoph Engel. Dictator games: A meta study. *Experimental economics*, 14:583–610, 2011.

[40] Owain Evans, Andreas Stuhlmüller, and Noah Goodman. Learning the preferences of ignorant, inconsistent agents. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2016.

[41] Yiding Feng, Chien-Ju Ho, and Wei Tang. Rationality-robust information design: Bayesian persuasion under quantal response. *arXiv preprint arXiv:2207.08253*, 2022.

[42] Emilio Ferrara. Should chatgpt be biased? challenges and risks of bias in large language models. *arXiv preprint arXiv:2304.03738*, 2023.

[43] Patrick J Fowler, Peter S Hovmand, Katherine E Marcal, and Sanmay Das. Solving homelessness from a complex systems perspective: insights for prevention responses. *Annual review of public health*, 40:465–486, 2019.

[44] Noufel Frikha, Stéphane Menozzi, et al. Concentration bounds for stochastic approximations. *Electronic Communications in Probability*, 17, 2012.

[45] Adrian Furnham and Hua Chu Boo. A literature review of the anchoring effect. *The journal of socio-economics*, 40(1):35–42, 2011.

[46] Yuan Gao, Sanmay Das, and Patrick Fowler. Homelessness service provision: a data science perspective. In *Workshops at the thirty-first AAAI conference on artificial intelligence*, 2017.

[47] Fabrizio Gilardi, Meysam Alizadeh, and Maël Kubli. Chatgpt outperforms crowd-workers for text-annotation tasks. *arXiv preprint arXiv:2303.15056*, 2023.

[48] Jean-Baptiste Gotteland, Nicolas Durand, Jean-Marc Alliot, and Erwan Page. Aircraft ground traffic optimization. In *ATM 2001, 4th USA/Europe Air Traffic Management Research and Development Seminar*, pages pp–xxxx, 2001.

[49] Ben Green and Yiling Chen. Disparate interactions: An algorithm-in-the-loop analysis of fairness in risk assessments. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, pages 90–99, 2019.

[50] Ben Green and Yiling Chen. The principles and limits of algorithm-in-the-loop decision making. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–24, 2019.

[51] Sachin Grover, Sailik Sengupta, Tathagata Chakraborti, Aditya Prasad Mishra, and Subbarao Kambhampati. Radar: Automated task planning for proactive decision support. *Human–Computer Interaction*, 35(5-6):387–412, 2020.

[52] Riccardo Guidotti, Anna Monreale, Salvatore Ruggieri, Franco Turini, Fosca Giannotti, and Dino Pedreschi. A survey of methods for explaining black box models. *ACM computing surveys (CSUR)*, 51(5):1–42, 2018.

[53] Moritz Hardt, Nimrod Megiddo, Christos Papadimitriou, and Mary Wootters. Strategic classification. In *Proceedings of the 2016 ACM conference on innovations in theoretical computer science*, pages 111–122, 2016.

[54] Chien-Ju Ho and Jennifer Wortman Vaughan. Online task assignment in crowdsourcing markets. In *26th AAAI Conference on Artificial Intelligence (AAAI)*, 2012.

[55] Chien-Ju Ho and Ming Yin. Working in pairs: Understanding the effects of worker interactions in crowdwork. *arXiv preprint arXiv:1810.09634*, 2018.

[56] Chien-Ju Ho, Yu Zhang, Jennifer Wortman Vaughan, and Mihaela van der Schaar. Towards social norm design for crowdsourcing markets. In *4th Human Computation Workshop (HCOMP)*, 2012.

[57] Chien-Ju Ho, Shahin Jabbari, and Jennifer Wortman Vaughan. Adaptive task assignment for crowdsourced classification. In *The 30th International Conference on Machine Learning (ICML)*, 2013.

[58] Chien-Ju Ho, Aleksandrs Slivkins, and Jennifer Wortman Vaughan. Adaptive contract design for crowdsourcing markets: Bandit algorithms for repeated principal-agent problems. In *15th ACM Conf. on Electronic Commerce (EC)*, 2014.

[59] Chien-Ju Ho, Aleksandrs Slivkins, Siddharth Suri, and Jennifer Wortman Vaughan. Incentivizing high quality crowdwork. In *Proceedings of the 24th International Conference on World Wide Web*, pages 419–429, 2015.

[60] Chien-Ju Ho, Rafael Frongillo, and Yiling Chen. Eliciting categorical data for optimal aggregation. In *30th Advances in Neural Information Processing Systems (NIPS)*, 2016.

[61] Chien-Ju Ho, Aleksandrs Slivkins, and Jennifer Wortman Vaughan. Adaptive contract design for crowdsourcing markets: Bandit algorithms for repeated principal-agent problems. *Journal of Artificial Intelligence Research*, 55:317 – 359, 2016.

[62] Kevin Anthony Hoff and Masooda Bashir. Trust in automation: Integrating empirical evidence on factors that influence trust. *Human factors*, 57(3):407–434, 2015.

[63] Zhiyu Huang, Haochen Liu, Jingda Wu, and Chen Lv. Conditional predictive behavior planning with inverse reinforcement learning for human-like autonomous driving. *IEEE Transactions on Intelligent Transportation Systems*, 2023.

[64] Yasuo Ishihara and Steve Johnson. Aircraft systems and methods for managing runway awareness and advisory system (raas) callouts, February 12 2019. US Patent 10,204,523.

[65] Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore. Reinforcement learning: A survey. *Journal of artificial intelligence research*, 4:237–285, 1996.

[66] Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decisions under risk. *Econometrica*, pages 263–291, 1979.

[67] Daniel Kahneman, Stewart Paul Slovic, Paul Slovic, and Amos Tversky. *Judgment under uncertainty: Heuristics and biases*. Cambridge university press, 1982.

[68] Daniel Kahneman, Jack L Knetsch, and Richard H Thaler. The endowment effect, loss aversion, and status quo bias. *The Journal of Economic Perspectives*, 5(1):193–206, 1991.

[69] Niklas Karlsson, George Loewenstein, and Duane Seppi. The ostrich effect: Selective attention to information. *Journal of Risk and uncertainty*, 38:95–115, 2009.

[70] Valmeekam Karthik, Sarath Sreedharan, Sailik Sengupta, and Subbarao Kambhampati. Radar-x: An interactive interface pairing contrastive explanations with revised plan suggestions. In *AAAI*, pages 16051–16053, 2021.

[71] Harmanpreet Kaur, Harsha Nori, Samuel Jenkins, Rich Caruana, Hanna Wallach, and Jennifer Wortman Vaughan. Interpreting interpretability: Understanding data scientists' use of interpretability tools for machine learning. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2020.

[72] Joshua Klayman and Young-Won Ha. Confirmation, disconfirmation, and information in hypothesis testing. *Psychological review*, 94(2):211, 1987.

[73] Jon Kleinberg and Manish Raghavan. How do classifiers induce agents to invest effort strategically? In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 825–844, 2019.

[74] Amanda Kube, Sanmay Das, and Patrick J Fowler. Allocating interventions based on predicted outcomes: A case study on homelessness services. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 622–629, 2019.

[75] Amanda Kube, Sanmay Das, Patrick J Fowler, and Yevgeniy Vorobeychik. Just resource allocation? how algorithmic predictions and human notions of justice interact. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 1184–1242, 2022.

[76] Amanda R Kube, Sanmay Das, and Patrick J Fowler. Community-and data-driven homelessness prevention and service delivery: optimizing for equity. *Journal of the American Medical Informatics Association*, 30(6):1032–1041, 2023.

[77] Tze Leung Lai and Herbert Robbins. Asymptotically efficient adaptive allocations rules. *Advances in Applied Mathematics*, 6:4–22, 1985.

[78] Vivian Lai and Chenhao Tan. On human predictions with explanations and predictions of machine learning models: A case study on deception detection. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, pages 29–38, 2019.

[79] Vivian Lai, Han Liu, and Chenhao Tan. " why is' chicago'deceptive?" towards building model-driven tutorials for humans. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2020.

[80] Wouter Kool Lauren Treiman, Chien-Ju Ho. Humans forgo reward to instill fairness into AI. Working paper, 2023.

[81] Zhuoshu Li, Kelsey Lieberman, William Macke, Sofia Carrillo, Chien-Ju Ho, Jason Wellen, and Sanmay Das. Incorporating compatible pairs in kidney exchange: A dynamic weighted matching model. In *Proceedings of the 2019 ACM Conference on Economics and Computation (EC)*, 2019.

[82] Q Vera Liao and Jennifer Wortman Vaughan. Ai transparency in the age of llms: A human-centered research roadmap. *arXiv preprint arXiv:2306.01941*, 2023.

[83] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.

[84] Han Liu, Vivian Lai, and Chenhao Tan. Understanding the effect of out-of-distribution examples and interactive explanations on human-ai decision making. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2):1–45, 2021.

[85] Yang Liu and Chien-Ju Ho. Incentivizing high quality user contributions: New arm generation in bandit learning. In *32nd AAAI Conference on Artificial Intelligence (AAAI)*, 2018.

[86] Jennifer M Logg, Julia A Minson, and Don A Moore. Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151:90–103, 2019.

[87] Zhuoran Lu and Ming Yin. Human reliance on machine learning models when performance feedback is limited: Heuristics and risks. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–16, 2021.

[88] Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems*, 2017.

[89] Anirudha Majumdar, Sumeet Singh, Ajay Mandlekar, and Marco Pavone. Risk-sensitive inverse reinforcement learning via coherent risk models. In *Robotics: science and systems*, volume 16, page 117, 2017.

[90] Vinayak Mathur, Yannis Stavrakas, and Sanjay Singh. Intelligence analysis of tay twitter bot. In *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)*, pages 231–236. IEEE, 2016.

[91] Daniel McFadden. Econometric models of probabilistic choice. *Structural analysis of discrete data with econometric applications*, 198272, 1981.

[92] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*, 2013.

[93] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529, 2015.

[94] Ray Moynihan, Jenny Doust, and David Henry. Preventing overdiagnosis: how to stop harming the healthy. *Bmj*, 344:e3502, 2012.

[95] Saumik Narayanan, Kassa Korley, Chien-Ju Ho, and Siddhartha Sen. Improving the strength of human-like models in chess. In *Human in the Loop Learning (HiLL) Workshop at NeurIPS*, 2022.

[96] Saumik Narayanan, Guanghui Yu, Wei Tang, Chien-Ju Ho, and Ming Yin. How does predictive information affect human ethical preferences? In *ACM Conference on AI, Ethics, and Society*, 2022.

[97] Saumik Narayanan, Guanghui Yu, Chien-Ju Ho, and Ming Yin. How does value similarity affect human reliance in ai-assisted ethical decision making? In *ACM Conference on AI, Ethics, and Society*, 2023.

[98] Nagarajan Natarajan, Inderjit S Dhillon, Pradeep K Ravikumar, and Ambuj Tewari. Learning with noisy labels. *Advances in neural information processing systems*, 26, 2013.

[99] Andrew Y Ng, Stuart J Russell, et al. Algorithms for inverse reinforcement learning. In *International Conference on Machine Learning*, 2000.

[100] Van Nguyen and S Tran. Conditional updates of answer set programming and its application in explainable planning. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, 2020.

[101] Van Nguyen, Stylianos Loukas Vasileiou, Tran Cao Son, and William Yeoh. Explainable planning using answer set programming. In *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning*, pages 662–666, 2020.

[102] Raymond S Nickerson. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, 2(2):175–220, 1998.

[103] Martin A Nowak, Karen M Page, and Karl Sigmund. Fairness versus reason in the ultimatum game. *Science*, 289(5485):1773–1775, 2000.

[104] Shutai Okamura, Takeshi Hatakeyama, Takahiro Yamaguchi, and Tsutomu Uenoyama. Radar detection system and radar detection method, January 19 2021. US Patent 10,895,638.

[105] Eli Pariser. *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin, 2011.

[106] Joseph T Pesik and David Matty. Determination of collision risks between a taxiing aircraft and objects external to the taxiing aircraft, February 4 2020. US Patent 10,553,123.

[107] Joshua C Peterson, David D Bourgin, Mayank Agrawal, Daniel Reichman, and Thomas L Griffiths. Using large-scale experiments and machine learning to discover theories of human decision-making. *Science*, 372(6547):1209–1214, 2021.

[108] Aleksandra Piktus, Christopher Akiki, Paulo Villegas, Hugo Laurençon, Gérard Dupont, Alexandra Sasha Luccioni, Yacine Jernite, and Anna Rogers. The roots search tool: Data transparency for llms. *arXiv preprint arXiv:2302.14035*, 2023.

[109] Forough Poursabzi-Sangdeh, Daniel G Goldstein, Jake M Hofman, Jennifer Wortman Wortman Vaughan, and Hanna Wallach. Manipulating and measuring model interpretability. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–52, 2021.

[110] Walter Quattrociocchi, Antonio Scala, and Cass R Sunstein. Echo chambers on facebook. *Available at SSRN 2795110*, 2016.

[111] Matthew Rabin and Ted O'Donoghue. Doing It Now or Later. *American Economic Review*, 1999.

[112] Jack W Rae, Sebastian Borgeaud, Trevor Cai, Katie Millican, Jordan Hoffmann, Francis Song, John Aslanides, Sarah Henderson, Roman Ring, Susannah Young, et al. Scaling language models: Methods, analysis & insights from training gopher. *arXiv preprint arXiv:2112.11446*, 2021.

[113] Manish Raghavan, Aleksandrs Slivkins, Jennifer Vaughan Wortman, and Zhiwei Steven Wu. The externalities of exploration and how data diversity helps exploitation. In *Conference on Learning Theory*, pages 1724–1738. PMLR, 2018.

[114] Deepak Ramachandran and Eyal Amir. Bayesian inverse reinforcement learning. In *IJCAI*, volume 7, pages 2586–2591, 2007.

[115] Kamalini Ramdas, Khaled Saleh, Steven Stern, and Haiyan Liu. Variety and experience: Learning and forgetting in the use of surgical devices. *Management Science*, 64(6):2590–2608, 2018.

[116] Gonzalo Ramos, Jina Suh, Soroush Ghorashi, Christopher Meek, Richard Banks, Saleema Amershi, Rebecca Fiebrink, Alison Smith-Renner, and Gagan Bansal. Emerging perspectives in human-centered machine learning. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–8, 2019.

[117] Charvi Rastogi, Yunfeng Zhang, Dennis Wei, Kush R Varshney, Amit Dhurandhar, and Richard Tomsett. Deciding fast and slow: The role of cognitive biases in ai-assisted decision-making. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW1):1–22, 2022.

[118] Vikas Raykar, Shipeng Yu, Linda Zhao, Gerardo Valadez, Charles Florin, Luca Bogoni, and Linda Moy. Learning from crowds. *Journal of Machine Learning Research*, 11:1297–1322, 2010.

[119] Amy Rechkemmer and Ming Yin. When confidence meets accuracy: Exploring the effects of multiple performance indicators on trust in machine learning models. In *CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2022.

[120] Sid Reddy, Anca Dragan, and Sergey Levine. Where do you think you're going?: Inferring beliefs about dynamics from behavior. *Advances in Neural Information Processing Systems*, 31, 2018.

[121] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. " why should i trust you?" explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1135–1144, 2016.

[122] Herbert Robbins and Sutton Monro. A stochastic approximation method. *The annals of mathematical statistics*, pages 400–407, 1951.

[123] Matthew J Salganik, Peter Sheridan Dodds, and Duncan J Watts. Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311(5762):854 – 856, 2006.

[124] Kristin E Schaefer, Jessie YC Chen, James L Szalma, and Peter A Hancock. A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. *Human factors*, 58(3):377–400, 2016.

[125] Clayton Scott. A rate of convergence for mixture proportion estimation, with application to learning from noisy labels. In *Artificial Intelligence and Statistics*, pages 838–846. PMLR, 2015.

[126] Sailik Sengupta, Tathagata Chakraborti, Sarath Sreedharan, Satya Gautam Vadlamudi, and Subbarao Kambhampati. Radar-a proactive decision support system for human-in-the-loop planning. In *AAAI Fall Symposia*, pages 269–276, 2017.

[127] Kyriacos Shiarlis, Joao Messias, and Shimon Whiteson. Inverse reinforcement learning from failure. *International Foundation for Autonomous Agents and Multiagent Systems*, 2016.

[128] Marybeth Shinn, Andrew L Greer, Jay Bainbridge, Jonathan Kwon, and Sara Zuiderveen. Efficient targeting of homelessness prevention services for families. *American journal of public health*, 103 (S2):S324–S330, 2013.

[129] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, et al. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*, 362(6419):1140–1144, 2018.

[130] Abneesh Singla, Srinivas D Gonabal, Pradeep Huncha, Vedavyas Rallabandi, Jaibir Singh, Sunil Kumar KS, et al. System and method for monitoring compliance with air traffic control instructions, August 25 2020. US Patent 10,755,583.

[131] Kenneth A Small. A discrete choice model for ordered alternatives. *Econometrica: Journal of the Econometric Society*, pages 409–424, 1987.

[132] Hummy Song, Anita L Tucker, Karen L Murrell, and David R Vinson. Closing the productivity gap: Improving worker productivity through public relative performance feedback and validation of best practices. *Management Science*, 64(6):2628–2649, 2018.

[133] Petru Soviany, Radu Tudor Ionescu, Paolo Rota, and Nicu Sebe. Curriculum learning: A survey. *International Journal of Computer Vision*, 130(6):1526–1565, 2022.

[134] Sarath Sreedharan, Tathagata Chakraborti, and Subbarao Kambhampati. Handling model uncertainty and multiplicity in explanations via model reconciliation. In *ICAPS*, pages 518–526, 2018.

[135] Vasileiou Loukas Stylianos, Alessandro Previti, and William Yeoh. On exploiting hitting sets for model reconciliation. In *Proceedings of AAAI*, 2021.

[136] Zhe Sun, Cheng Zhang, Pingbo Tang, Yuhao Wang, and Yongming Liu. Bayesian network modeling of airport runway incursion occurring processes for predictive accident control. In *Advances in Informatics and Computing in Civil and Construction Engineering: Proceedings of the 35th CIB W78 2018 Conference: IT in Design, Construction, and Management*, pages 669–676. Springer, 2019.

[137] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. The MIT Press, second edition, 2018.

[138] Wei Tang and Chien-Ju Ho. Bandit learning with biased human feedback. In *Proceedings of the 18th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 1324–1332, 2019.

[139] Wei Tang and Chien-Ju Ho. On the bayesian rational assumption in information design. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, volume 9, pages 120–130, 2021.

[140] Wei Tang, Ming Yin, and Chien-Ju Ho. Leveraging peer communication to enhance crowdsourcing. In *The World Wide Web Conference*, pages 1794–1805. ACM, 2019.

[141] Wei Tang, Chien-Ju Ho, and Yang Liu. Differentially private contextual dynamic pricing. In *Proceedings of the 19th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2020.

[142] Wei Tang, Chien-Ju Ho, and Yang Liu. Bandit learning with delayed impact of actions. *Advances in Neural Information Processing Systems*, 34:26804–26817, 2021.

[143] Wei Tang, Chien-Ju Ho, and Yang Liu. Linear models are robust optimal under strategic behavior. In *International Conference on Artificial Intelligence and Statistics*, pages 2584–2592. PMLR, 2021.

[144] Kenneth E Train. *Discrete choice methods with simulation*. Cambridge university press, 2009.

[145] Long Tran-Thanh, Archie Chapman, Enrique Munoz De Cote, Alex Rogers, and Nicholas R Jennings. Epsilon–first policies for budget–limited multi-armed bandits. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 1211–1216, 2010.

[146] Amos Tversky and Daniel Kahneman. Belief in the law of small numbers. *Psychological bulletin*, 76 (2):105, 1971.

[147] Amos Tversky and Daniel Kahneman. Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science*, 185(4157):1124–1131, 1974.

[148] Amos Tversky and Daniel Kahneman. The framing of decisions and the psychology of choice. *Science*, 211(4481):453–458, 1981.

[149] Stylianos Loukas Vasileiou, Alessandro Previti, and William Yeoh. On exploiting hitting sets for model reconciliation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 6514–6521, 2021.

[150] Stylianos Loukas Vasileiou, William Yeoh, Tran Cao Son, and Alessandro Previti. Explanations as model reconciliation via probabilistic logical reasoning. In *Proceedings of the Explainable Logic-Based Knowledge Representation (XLoKR)*, 2021.

[151] Oleksandra Vereschak, Gilles Bailly, and Baptiste Caramiaux. How to evaluate trust in ai-assisted decision making? a survey of empirical methodologies. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2):1–39, 2021.

[152] John von Neumann and Oscar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

[153] Xin Wang, Yudong Chen, and Wenwu Zhu. A survey on curriculum learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(9):4555–4576, 2021.

[154] Christopher J.C.H. Watkins and Peter Dayan. Q-learning. *Machine learning*, 8(3-4):279–292, 1992.

[155] Christabel Wayllace, Sunwoo Ha, Yuchen Han, Jiaming Hu, Shayan Monadjemi, William Yeoh, and Alvitta Ottley. Dragon-v: detection and recognition of airplane goals with navigational visualization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 13642–13643, 2020.

[156] H Gilbert Welch, Lisa Schwartz, and Steve Woloshin. *Overdiagnosed: making people sick in the pursuit of health*. beacon press, 2012.

[157] Jacob Whitehill, Ting fan Wu, Jacob Bergsma, Javier R. Movellan, and Paul L. Ruvolo. Whose vote should count more: Optimal integration of labels from labelers of unknown expertise. In *Advances in Neural Information Processing Systems (NIPS)*, 2009.

[158] Marty J Wolf, K Miller, and Frances S Grodzinsky. Why we should have seen that coming: comments on microsoft's tay" experiment," and wider implications. *Acm Sigcas Computers and Society*, 47(3): 54–64, 2017.

[159] Michael Yeomans, Anuj Shah, Sendhil Mullainathan, and Jon Kleinberg. Making sense of recommendations. *Journal of Behavioral Decision Making*, 32(4):403–414, 2019.

[160] Ming Yin, Jennifer Wortman Vaughan, and Hanna Wallach. Understanding the effect of accuracy on trust in machine learning models. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 279. ACM, 2019.

[161] Guanghui Yu and Chien-Ju Ho. Environment design for biased decision makers. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2022.

[162] Kaiqing Zhang, Zhuoran Yang, and Tamer Başar. Multi-agent reinforcement learning: A selective overview of theories and algorithms. *Handbook of reinforcement learning and control*, pages 321–384, 2021.

[163] Yunfeng Zhang, Q Vera Liao, and Rachel KE Bellamy. Effect of confidence and explanation on accuracy and trust calibration in ai-assisted decision making. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pages 295–305, 2020.

[164] Jiangchuan Zheng, Siyuan Liu, and Lionel M Ni. Robust bayesian inverse reinforcement learning with sparse behavior noise. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2014.

[165] Yudian Zheng, Guoliang Li, Yuanbing Li, Caihua Shan, and Reynold Cheng. Truth inference in crowdsourcing: Is the problem solved? *Proceedings of the VLDB Endowment*, 2017.

[166] Meixin Zhu, Yinhai Wang, Ziyuan Pu, Jingyun Hu, Xuesong Wang, and Ruimin Ke. Safe, efficient, and comfortable velocity control based on reinforcement learning for autonomous driving. *Transportation Research Part C: Emerging Technologies*, 117:102662, 2020.

[167] Brian D Ziebart, Andrew Maas, J Andrew Bagnell, and Anind K Dey. Maximum entropy inverse reinforcement learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2008.