The Impacts of Algorithmic Work Assignment on Fairness Perceptions and Productivity: Evidence from Field Experiments

Bing Bai[1], Hengchen Dai[2], Dennis J. Zhang[1], Fuqiang Zhang[1], Haoyuan Hu[3]

1. Olin Business School, Washington University in St. Louis, St. Louis, MO, USA
2. Anderson School of Management, University of California, Los Angeles, Los Angeles, CA, USA
3. Alibaba Group, Hangzhou, China

Problem Definition: We study how algorithmic (vs. human-based) task assignment processes change task recipients' fairness perceptions and productivity.

Academic/Practical Relevance: Since algorithms are widely adopted by businesses and often require human involvement, understanding how humans perceive algorithms is instrumental to the success of algorithm design in operations. Particularly, the growing concern that algorithms may reproduce inequality historically exhibited by humans calls for research about how people perceive the fairness of algorithmic decision-making relative to traditional human-based decision-making and, consequently, adjust their work behaviors.

Methodology: In a 15-day-long field experiment with Alibaba Group in a warehouse where workers pick products following orders (or "pick lists"), we randomly assigned half of the workers to receive pick lists from a machine that ostensibly relied on an algorithm to distribute pick lists, and the other half to receive pick lists from a human distributor.

Results: Despite that we used the same underlying rule to assign pick lists in both groups, workers perceive the algorithmic (vs. human-based) assignment process as fairer by 0.94-1.02 standard deviations. This yields productivity benefits: receiving tasks from an algorithm (vs. a human) increases workers' picking efficiency by 17.35%-19.39%. The algorithmic assignment produces larger productivity gains among workers for whom perceived fairness has a stronger effect on productivity, including more educated workers and workers who care more about the difficulty of their pick lists. We replicate the main results in another field experiment and show via online experiments that people in the U.S. also view algorithmic task assignment as fairer.

Managerial Implications: We demonstrate that algorithms can have broader impacts beyond offering greater efficiency and accuracy than humans: introducing algorithmic assignment processes may enhance fairness perceptions and productivity. This insight can be utilized by managers and algorithm designers to better design and implement algorithm-based decision making in operations.

Key words: Behavioral Operations, Field Experiment, Productivity, Fairness, Artificial Intelligence

1. Introduction

With the increasing availability of data and the development of information technologies, companies are rapidly implementing algorithms to process a large amount of data in order to efficiently make daily operational decisions (McAfee and Brynjolfsson 2017). For example, digital service platforms such as Uber and Airbnb instantly match customers with service providers, taking high-dimensional information into account (e.g., customers' willingness to pay, service providers' availability) in their algorithms. Ad platforms such as Facebook and Google combine advertising algorithms with rich data about consumers to identify specific audience groups for which to display certain ads.

Such growing interest in using algorithms in practice has inspired a large stream of research dedicated to improving algorithms' performance (e.g., Kropp and Carlson 1984, Bhandari et al. 2008, Mookerjee et al. 2017, Zhang and Kulkarni 2018). However, in many domains of daily operations, algorithms rely on human involvement to complete tasks. For example, retailing platforms such as Alibaba use algorithms to determine which set of items should be packed into which box but need human workers in warehouses to pack the items according to algorithmic prescriptions (Sun et al. 2020). Similarly, sales platforms such as Salesforce use algorithms to decide which product to be advertised to whom but need human salespeople to make sales pitches to customers following algorithmic recommendations.

Thus, another fundamental question about algorithm development in operations management is how humans perceive and interact with algorithms. A growing body of work has begun to study this question from both operational and psychological perspectives (Dijkstra et al. 1998, Dietvorst et al. 2015, 2018, Leung et al. 2018, Castelo et al. 2019, Dietvorst and Bharti 2020, Jago 2019, Logg et al. 2019, Luo et al. 2019, Newman et al. 2020, Sun et al. 2020, Xu and Jago 2020). This literature has largely highlighted that people are reluctant to use algorithms and prefer instead to defer to judgments made by a human, regardless of whether the human is a peer, an expert, or their own self. This reluctance may originate from people's need for control (Dietvorst et al. 2018), need for identity signaling (Leung et al. 2018), or their negative impressions about algorithms (such as lack of authenticity, less empathy, and lower competence in certain contexts; Dietvorst and Bharti 2020, Jago 2019, Luo et al. 2019, Newman et al. 2020).

In particular, one growing concern is that algorithms may produce or reproduce discriminatory outcomes and lead to new or more systematic biases than what humans have historically exhibited. Critics are concerned that algorithms may reproduce disparities across demographic groups due to (unconscious or conscious) biases in the objective functions that algorithms are set up to optimize or in the data used to train algorithms (see Cowgill and Tucker 2019 a review). Scholars sharing this concern have provided empirical support for the existence of algorithmic bias in the domains of judicial decision-making (Angwin et al. 2016), hiring (Datta et al. 2015), targeted advertising (Lambrecht and Tucker 2019), and health care (Obermeyer et al. 2019). Motivated by this concern, researchers have studied various ways of defining and enforcing fairness when designing algorithms (Corbett-Davies et al. 2017, Kleinberg et al. 2018). Extending this line of work on how and why algorithms actually produce unfair judgment, we study how workers *perceive* the fairness of algorithmic decisions and how such fairness perceptions affect their behavior when algorithms are used to make decisions related to workers' tasks.

In work settings, algorithms are increasingly replacing human decision makers to determine allocations of resources and tasks (e.g., delivery trips, customers, cases) across employees. We specifically examine how an algorithmic task assignment process, relative to a human-based task assignment process, changes task recipients' fairness perceptions as well as its implications for their productivity. To causally answer these questions, we conducted field experiments in collaboration with Alibaba—the largest retailing platform in China—in its warehouse setting. In recent years, e-commerce warehouses have started digitizing equipment and applying algorithms to many key tasks within warehouses, such as picking, routing, scheduling, and bin packing (Sun et al. 2020). We focus on picking tasks for which workers receive a picking order (denoted as a "pick list") and follow the pick list to pick products from different stocking shelves.

We conducted two field experiments in a warehouse. Before and in between our two experiments, hard-copy pick lists were periodically printed out and placed on a table at the distribution station for workers to take at their discretion. During our experiments, workers were randomly assigned to one of two groups: workers in the algorithm group received picking tasks from a machine that ostensibly relied on an algorithm to distribute pick lists, whereas workers in the human group received picking tasks from a human distributor. In order to cleanly identify the impact of workers' perceptions of algorithmic (vs. human-based) assignment process on productivity, we removed other differences between these two types of task assignment processes. Specifically, unbeknownst to any worker, the human distributor and the algorithm actually assigned tasks from the same pool of available pick lists using similar rules at any given point, which guaranteed that the characteristics of the pick lists as well as the matches between pick lists and workers were comparable between two groups. Also, workers in both groups could only receive their tasks from a central distribution station, which allowed us to control how long workers needed to walk to get their next task. Moreover, we deliberately did not give workers any information about how the algorithm assigned the tasks so that workers were not biased against or in favor of specific algorithms. In essence, we kept the objective nature of these two task assignment processes as similar as we could, so the difference we observe between these two assignment processes can be attributed to workers' beliefs and perceptions about the differences between algorithmic and human-based assignment.

The first, main experiment involved 50 temporary workers for 15 days in August-September, 2019. We collected data about all $4,486$ pick lists completed by workers during this experiment, along with 108 daily questionnaires from them. We present four key findings. First, we find that workers hold different views about the fairness of these two assignment processes: workers receiving tasks from an algorithm on average perceive their assignment process as more fair than workers receiving tasks from a human distributor, and the difference is 0.94-1.02 standard deviations (depending on our model specifications). Second, we document productivity differences between these two assignment processes: receiving tasks from an algorithm significantly increases workers' picking efficiency by 17.35%-19.39%, compared to receiving tasks from a human distributor. Third, we estimate via an instrumental variable approach that a one-standard-deviation increase in fairness perceptions can boost workers' picking efficiency by 16.04%-16.98%. Finally, we find that the algorithmic assignment process produces a larger productivity gain for workers who are more (vs. less) emotionally sensitive to the difficulty of their tasks and for workers with a higher (vs. lower) education level, because fairness perceptions play a bigger role in driving productivity for these subsamples of workers. We conducted the second field experiment with a nonoverlapping sample of workers in December, 2019-January, 2020 and validated the robustness of our main results.

To further validate our findings from the field, we also conducted an online experiment to study the effect of algorithmic (vs. human-based) task assignment on perceived fairness among a different population—201 people in the United States recruited from an online labor market (Amazon's Mechanical Turk). Study participants imagined working in a warehouse and receiving picking tasks from either a machine or a human distributor. Despite imagining receiving the same picking tasks, people on average perceived the assignment process run by a machine based on an algorithm as more fair than the process run by a human. We replicated this pattern in another online experiment with a slightly different design.

In summary, we empirically examine people's psychological and behavioral responses to algorithmic decision-making processes across experiments in different settings, and we provide the first field experiments in an actual workplace to study this issue. By keeping the nature of pick lists the same, our design provides a clean and conservative test of how people perceive algorithmic (vs. human-based) decision-making processes and how people behave after receiving decisions made by these processes. Theoretically, this angle differentiates our study from the large body of research that examines sources of algorithm-engendered biases and compares algorithms and humans in the actual level of inequality and discrimination they produce (Dwork et al. 2012, Angwin et al. 2016, Kleinberg et al. 2016, Caliskan et al. 2017, Chouldechova 2017, Cowgill 2018, Kleinberg et al. 2018, Lambrecht and Tucker 2019, Obermeyer et al. 2019). Practically, through this unique design, our findings can help companies understand workers' perceptions about algorithmic decision-making processes and optimize the framing of task assignment processes. Altogether, our research complements the existing literature about human and algorithm collaboration, highlights the importance of understanding workers' fairness perceptions about work assignment when utilizing algorithms, and provides insights for designing better human-algorithm collaboration in daily operations.

The rest of the paper is organized as follows. Section 2 reviews the relevant literature and our theoretical contributions. Section 3 develops our hypotheses. In Section 4, we introduce our field setting and experimental design. Sections 5 and 6 present the effects of algorithmic assignment on fairness perceptions and productivity as well as explore how these effects vary by workers. Section 7 reports an online experiment testing the effects of algorithmic assignment on fairness perceptions. We discuss implications of our findings in Section 8. Our replication field and online experiments are reported in Online Appendices.

2. Literature Review and Theoretical Contributions

Our work is mainly related to four research areas: human collaboration with algorithms, algorithmic bias, operations management research about automation, and behavioral operations.

2.1. Human Collaboration with Algorithms

Our work is closely connected to the growing stream of literature studying how people perceive and react to algorithms and automation. The primary focus of this literature has been on examining whether humans, as *users* of algorithms, are willing to rely on algorithmic prescriptions and utilize automated systems. With a few exceptions (Dijkstra et al. 1998, Logg et al. 2019), research in this area has largely documented *algorithm aversion*, whereby people are reluctant to utilize algorithms and automation (compared to their own judgment, human experts' advice, or their peers' aid), despite the fact that algorithms give identical output or, in some cases, even superior performance than humans (Dietvorst et al. 2015, 2018, Leung et al. 2018, Dietvorst and Bharti 2020, Jago 2019, Longoni et al. 2019, Luo et al. 2019). This may happen because people have less error tolerance for algorithms than for human judges (Dietvorst et al. 2015), want to exert control and signal their social identity by actively making or influencing a certain decision (Dietvorst et al. 2018, Leung et al. 2018), or (mistakenly) believe that algorithms are less competent than humans in certain circumstances (e.g., when making forecasts in inherently uncertain domains, when expressing emotional connection and authenticity, and when performing a subjective task; Dietvorst and Bharti 2020, Castelo et al. 2019, Jago 2019, Luo et al. 2019).

More recently, this literature has begun to examine how people as *recipients* of decisions made about them (e.g., employees who receive personnel decisions, students who receive housing allocation decisions) respond to algorithmic versus human-based decision processes (Longoni et al. 2019, Newman et al. 2020, Xu and Jago 2020). This line of research so far has found that people view algorithms as less capable of taking into account their unique, contextual, and personal characteristics (Longoni et al. 2019, Newman et al. 2020); as a result, people perceive algorithmic (vs. human-based) decision-making as less procedurally fair and express less commitment to organizations (e.g., their companies, schools) that use algorithms (rather than humans) for decision making (Newman et al. 2020, Xu and Jago 2020).

We make several contributions to this literature. First, while recent research suggests that people disfavor algorithms when they want decision-making processes to consider their unique and personal characteristics (Castelo et al. 2019, Longoni et al. 2019, Newman et al. 2020, Xu and Jago 2020), we recognize that people often have the equality motive—that is, they would like to receive equal treatment and opportunity relative to others (Dawes et al. 2007, Rai and Fiske 2011). The equality motive is a universal motive and often occurs in business operations (Rai and Fiske 2011). We complement prior research by documenting that algorithmic decision-making processes are viewed more favorably than human-based decision-making processes in settings where people prioritize the equality motive over other motives that highlight uniqueness and consideration of personal characteristics. Second, while prior research has focused on how people collaborate with algorithms on prediction tasks and consumer decision-making, we examine how employees perceive algorithms that determine their tasks at work. Our empirical context in the field studies—a labor-intensive working environment—is also a complement to the literature. Third, while the prior research reviewed above has largely used laboratory and online experiments, we conducted field experiments in a common operation setting (warehouse operations) to provide more external validity of our insights. Fourth, going beyond examining people's *perceptions* of algorithms that determine their outcomes, we further study employees' work behaviors and find a downstream consequence of algorithmic work assignment process on productivity, which has been overlooked by prior literature.

2.2. Algorithmic Bias

Our paper is also related to the emerging literature studying biases and discrimination engendered by algorithms. Scholars are concerned that algorithms may inadvertently reproduce, codify, or even amplify disparities due to biases in objective functions, people who build the algorithms, or historical data (Cowgill and Tucker 2019), and have provided evidence that algorithms perpetuate existing inequality in a variety of domains (Dwork et al. 2012, Datta et al. 2015, Angwin et al. 2016, Kleinberg et al. 2016, Caliskan et al. 2017, Chouldechova 2017, Kleinberg et al. 2018, Lambrecht and Tucker 2019, Obermeyer et al. 2019). This concern has motivated researchers to study how to define and enforce fairness when designing algorithms (Corbett-Davies et al. 2017, Kleinberg et al. 2018). Despite the concern around algorithmic bias, the research that compares algorithms to human decision makers, although scarce, suggests that algorithmic judgment appears less biased than human judgement, even when algorithms are trained on historical data involving biased human decisions (Kleinberg et al. 2017, Cowgill 2018). This provides some empirical support for the more positive view that the use of AI could have positive implications for social equality and fairness by taking biased humans out of the equation.

While prior work in this literature has focused on identifying when algorithms produce biased outcomes and comparing algorithms to humans in the actual level of fairness generated, we study people's *perceptions* of algorithms' ability to deliver fair treatments. We ask the fundamental question of how knowing that one's outcome is determined by an algorithm (vs. a human) affects people's perceived fairness about the decision process, which subsequently influences their behaviors. To examine this question, we keep the underlying decision-making logic and the assigned outcomes the same but investigate how people's perceptions change when they are led to believe that their outcomes are decided by an algorithm rather than a human decision maker.

2.3. Automation in Operations Management

Our research adds to a large body of literature in operations management studying problems that arise in the presence of automation, particularly research that incorporates the role of humans in the design of automated systems (Van Donselaar et al. 2010, Ball and Ghysels 2017, Karacaoglu et al. 2018, Zhang and Kulkarni 2018, Choudhary et al. 2020, Li et al. 2020, Sun et al. 2020). For example, Van Donselaar et al. (2010) study how managers' deviation from advice given by an automated system could yield insights for how to improve the efficiency of the automated system. Relatedly, Sun et al. (2020) find that workers deviate from algorithmic prescriptions because they sometimes have information and behavioral biases that algorithms have not considered. Sun et al. (2020) further find that using a machine learning approach to incorporate these drivers of deviations into algorithm design can improve operational efficiency. While prior research in this area has focused on how to make algorithms and automated systems more powerful (e.g., by learning from humans' deviations), we study how people's perceptions about automation affect their efficiency. We show, via field experiments, that in the presence of automation, psychological factors such as fairness perceptions impact worker productivity.

2.4. Behavioral Operations

Finally, our work builds on the behavioral operations literature. This literature has documented a number of behavioral and psychological drivers of productivity, such as team familiarity, time pressure, peer pressure, quality monitoring, and free-rider effect (e.g., Huckman et al. 2009, Aksin et al. 2015, Ibanez and Toffel 2020, Tan and Netessine 2019, Wu and Wang 2019, Xu and Zhu 2020), mostly based on archival data analysis. Through longitudinal field experiments, we document that perceived fairness about task assignment is another important driver of productivity. Also, the behavioral operations literature has shown that people fall prey to behavioral biases in many operations settings, such as framing (Buell and Norton 2011, Kc 2020). For example, Kc (2020) shows that the framing of patients' admission time may affect doctors' discharge decisions. Our finding that people view a decision process as fairer and work more productively when the process is seemingly driven by an algorithm (vs. a human) provides an example where the framing of work assignment affects operational efficiency.

3. Hypothesis Development

In this section, we present two hypotheses regarding how assigning tasks via algorithms versus humans affects task recipients' fairness perceptions and productivity in prevalent work settings where workers tend to prioritize the equality motive. Key determinants of people's perceived fairness of a process used to make and implement allocation decisions include whether the process is free from decision makers' personal biases, applies decision rules consistently across people and across time, and uses appropriate factual information to make decisions (Leventhal et al. 1980, Tyler 1989, Colquitt et al. 2001). People may worry that a human decision maker would consciously or unconsciously exhibit bias in favor of some individuals for unjustifiable reasons (e.g., close relationships, physical attractiveness), but they may expect algorithms to be free of these personal biases and more capable of consistently applying rules and producing equal outcomes across individuals. Thus, we hypothesize that:

Hypothesis 1. *Workers perceive a task assignment process as more fair if they believe the process is implemented by an algorithm than if they believe the process is implemented by a human.*

Our next hypothesis pertains to how algorithmic (vs. human-based) assignment affects productivity. Research in psychology, organizational behavior, and behavioral economics consistently suggests that people desire fair treatments and behave differently at work in accordance to whether they think they are fairly treated in their organizations (see Cohen-Charash and Spector 2001, Colquitt et al. 2001, Fehr et al. 2009, Greenberg and Colquitt 2013 for reviews of relevant research). In particular, meta-analyses of hundreds of studies suggest that procedural fairness perceptions have a moderately positive correlation with work performance on average ($r = 0.30$; Colquitt et al. 2001) and that the relationship is stronger among actual employees in work settings ($r = 0.47$) as opposed to students in laboratory studies (Cohen-Charash and Spector 2001). Building on prior research, we predict that in our research settings, as an algorithmic task assignment process increases people's perceived fairness, it should subsequently have a positive impact on their productivity.

Hypothesis 2. *Workers are more productive if they believe their task assignment process is implemented by an algorithm than if they believe the process is implemented by a human.*

4. Experiment Design and Data

4.1. Field Setting and Experiment Design

Our field experiments were conducted in collaboration with Alibaba. In 2013, along with five package delivery companies, Alibaba co-founded Cainiao Network (hereafter, "Cainiao"), a logistic platform operator dedicated to digitizing the shipping industry and building a smart logistic network nationally and globally. Cainiao has the largest bonded warehouse network in China and manages more than 60% packages from Alibaba's Chinese retail marketplaces. Its vision is to achieve 24-hour delivery anywhere in China and 72-hour delivery anywhere globally. (See https://www.alibabagroup.com/en/ir/presentations/Investor_Day_2019_CainiaoNetwork.pdf.)

We study one core task that workers perform in warehouses: picking, which requires that workers pick certain products from different shelves following specific picking orders (or "pick lists"). When an online purchase order is placed on Alibaba, Cainiao's warehouse management system first decides which warehouse should fulfill the order based on the Stock Keeping Units (SKUs) included in the order and their stocking information. After accumulating a number of purchase orders for a given warehouse, the system generates a list of pick lists for this warehouse, with each pick list usually covering multiple purchase orders. A pick list contains information about products that a worker should pick, including SKU name, the quantity the worker should pick for each SKU, and the stocking location of each SKU (see Online Appendix A for an example pick list.)

We conducted two experiments in one of Cainiao's warehouses where picking workers were paid hourly. Before and in between the experiments, a staff member in the warehouse periodically printed out pick lists as physical hard copies and placed them on a table at the pick list distribution station so that workers could get pick lists themselves based on their own preferences. During our experiments, we manipulated how picking workers received pick lists and randomly assigned workers into either the human group or the algorithm group. We set up two tables side by side at the distribution station, one for the human group and the other for the algorithm group.

In the human group, hard-copy pick lists were printed and assigned by a human distributor. Specifically, a human distributor stood at the human-based assignment table and periodically printed out a stack of hard-copy pick lists from the pool of available pick lists. The selection of pick lists from the pool and the printing order were designed to be random. When a worker in the human group came to the human-based assignment table, the human distributor handed the worker a hard-copy pick list. Upon receiving the pick list, the worker scanned the bar code on the pick list using a radio-frequency hand-held monitor, at which time Cainiao's system would record the starting time of this pick list. After scanning the bar code on the hard-copy pick list, the picking worker no longer needed the hard-copy pick list since she could access the pick list's information on her hand-held monitor. In the algorithm group, picking tasks were assigned to workers by an algorithm. Specifically, when a worker in the algorithm group came to the distribution station for her next pick list, she would scan a bar code on the algorithmic assignment table at the station. This would trigger the algorithm to randomly choose a pick list from the pool of available pick lists and then display the selected pick list on the worker's hand-held monitor. At that time Cainiao's system would record the starting time of this pick list.

From this point on, the process of completing a pick list was the same in the algorithm and human groups. Upon the pick list showing up on their monitor, workers in both groups would walk to the stocking location of the first SKU on their pick list. Once they found the first SKU and put the corresponding quantity into a cart, they scanned the bar code of the first SKU to record that they successfully picked the first SKU. Then information about the next SKU would show up on the hand-held monitor. When workers picked the last SKU in their pick list and scanned its bar code, the pick list would be marked as completed and the finish time would be recorded. Figure 1 illustrates the pick list generation and picking process for both groups of workers.

Note that the pick list assignment process in both the human and algorithm conditions differed from how pick lists were assigned in this warehouse before and in between our experiments. Thus, the preexperiment assignment process could not have served as an anchor that differently affected workers' perceptions of algorithmic versus human-based assignment process during our experiments. Workers and human distributors were unaware of the objectives of our experiments or our hypotheses, and they did not have information about the algorithm in use. (It is common in China for workers in labor-intensive jobs (e.g., factories, warehouses) to simply complete tasks as instructed without asking about why certain procedures are implemented.) We also did not disclose our research objectives or hypotheses to the operation manager in the warehouse.

During our conversations with workers in other warehouses that had humans distribute hard-copy pick lists (rather than let workers take pick lists laid on a table at their own discretion, as in our collaborating warehouse before our experiments), a few workers expressed concerns that human distributors might assign easier tasks to workers they were familiar with and that factors such as appearance might play a role in human distributors' allocation decisions. To cleanly examine how workers *perceive* algorithmic assignment processes (relative to human-based assignment processes), we need to make sure that the accessibility and distribution of pick lists are not statistically different between those two conditions so that the two conditions only differ in workers' perceived distributor (i.e., algorithm vs. human). We took several measures to ensure this.

First, pick lists assigned to workers at any given point were drawn from the same pool of pick lists using the same underlying rule, regardless of whether workers received pick lists from a human distributor or an algorithm. Specifically, in the human group, as we mentioned above, when human distributors periodically printed out a stack of hard-copy pick lists, pick lists were randomly selected from the pool of available pick lists and printed out in a random order. During our experiments, we instructed human distributors in our collaborating warehouse to give out pick lists in the same (random) order in which pick lists were printed. This instruction prevented human distributors from handing out pick lists at their own discretion (as they might have done without our explicit instructions). In the algorithm group, the algorithm by design randomly selected a pick list from the pool of available pick lists each time. Therefore, in essence, workers in both groups received pick lists that were randomly drawn from a common pool of pick lists. Our following randomization check in Section 4.3 also confirms that the characteristics of pick lists are comparable between conditions. Second, during our experiments, workers in both groups had to walk to the same location to obtain their next pick list, which allowed us to eliminate the effects of different walking distances on productivity. In other words, it cannot be the case that workers in the algorithm group walked less (or more), were less (or more) fatigued, and thus were more (or less) productive than workers in the human group. Third, we sought to make the process of receiving pick lists equally simple for workers in both groups, so differences in productivity could not be driven by how inconvenient workers found the assignment process. Indeed, as shown in 4.3, workers in both groups rated the process of receiving pick lists as similarly convenient.

A potential concern about our experimental design is interference between workers; that is, the behavior of a particular worker may depend not only on her own pick list assignment process but also on the assignment process experienced by others in the warehouse (e.g., because they may communicate with each other). Following Aronow (2012), we conducted an ex-post statistical test to confirm that such interference between workers was unlikely to drive our results (See detail in Online Appendix B). While other experimental designs may avoid such potential interference between workers, we consider our approach the cleanest among all feasible approaches (similar to Fryer Jr et al. 2012, Hossain and List 2012, Bloom et al. 2015). One alternative design is to randomly assign different days into one of the two conditions (e.g., Buell et al. 2017). Though this approach avoids co-location of the algorithm and human groups, it has two significant limitations in our setting. First, workers who work in a warehouse for multiple days would likely experience both the algorithm and human conditions in this alternative design, in which case interference between conditions could arise because workers' experiences with one condition may affect their perceptions and behavior in the other condition. Second, in our setting, picking tasks differ substantially across work days (e.g., promotion days versus non-promotion days). Considering the frequency of promotions for various product categories at Alibaba, we need to run our experiments for at least a few months to get a comparable set of work days assigned to the algorithm group versus the human group, which is infeasible in our setting. Another alternative design is to run the experiment in many warehouses and assign workers to the human or algorithm condition at the warehouse level. Though this would ensure that workers in one condition do not communicate with workers in the other condition, this approach has its own challenge. Picking tasks usually differ greatly across warehouses in terms of the number of items in a pick list, the number of stocking areas covered by a pick list, and walking distance, and these factors could largely affect workers' productivity. Ideally, this problem would be resolved if we could run our experiments in hundreds of warehouses (as in Duflo and Saez 2003,

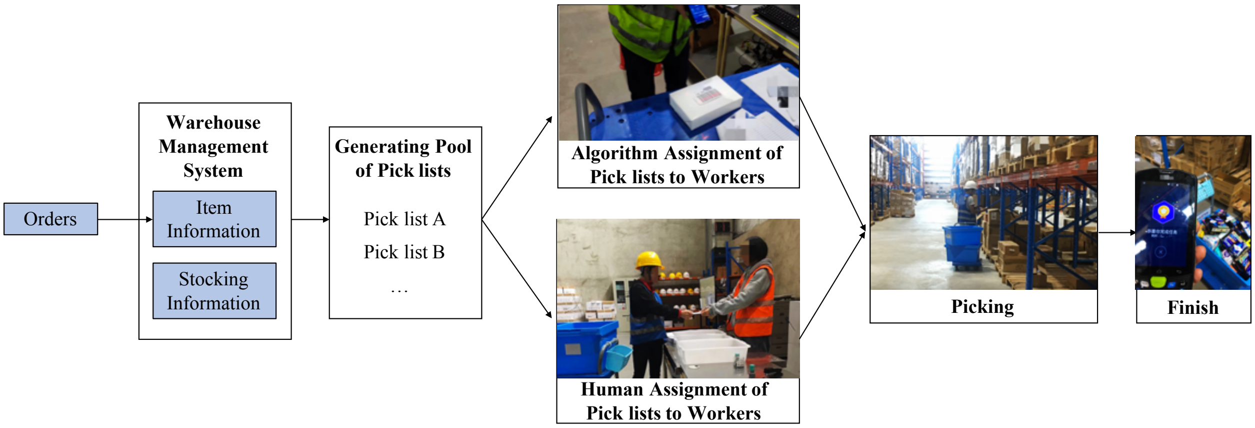


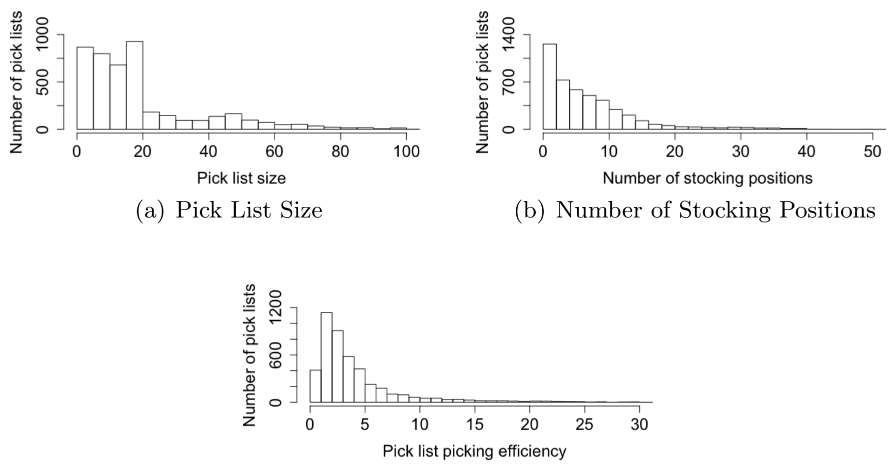
Figure 1 **Flow Chart of Picking Process for Both Groups**

Banerjee et al. 2007); however, this is logistically infeasible based on Cainiao's current priorities and the logistics of setting up the experiment in a standardized way across a large number of warehouses.

Our two field experiments had the same design but were run at two different times involving nonoverlapping samples of workers. We focus on the first field experiment in this paper, and report the second field experiment as a replication study in Online Appendix C. Our first, main experiment spanned 15 days, starting from August 20, 2019 and ending on September 6, 2019. On August 25-26 and August 30, 2019, the warehouse had a much heavier workload than usual due to Alibaba's platform-wide "shopping holidays" so the experiment was temporally halted on these days. This field experiment involved 50 temporary workers. On average, they worked 2.16 days during our experiment, yielding a total of 108 person-day pairs. These workers completed 4,486 pick lists in total.

4.2. Survey Design and Data

During our experiment, we collected two types of data: (1) operations data from the warehouse management system tracking the characteristics and processing time of each pick list, and (2) workers' responses to surveys that we administered every day. For each pick list, we track three characteristics that capture key information visible on a pick list (as shown in Online Appendix A)—the pick list size (i.e., the total quantity of items to be picked in the pick list), the inventory area (i.e., one of the two regions in the warehouse where items in the pick list were stocked), and the number of stocking positions (i.e., the number of shelf positions in which items in the pick list were stocked). In addition, we track the identifier of the worker who handled this pick list and the times when the worker started versus completed the pick list. Figure 2 provides the distributions of pick list statistics across 4,486 pick list observations in our first experiment. Figure 2(a) shows the distribution of pick list size. Since the picking carts that workers used to temporarily store



(a) Pick List Size (b) Number of Stocking Positions

(c) Picking Efficiency

Figure 2 Distributions of Pick List Statistics in the Field Experiment

products they collected had capacity limits, most (72.96%) pick lists contained no more than 20 items. Figure 2(b) shows the distribution of the number of stocking positions. Since pick lists were intended to combine items in the same stocking position, the number of stocking positions was usually smaller than the number of items in a pick list. Figure 2(c) displays the distribution of picking efficiency across pick lists. Picking efficiency, defined as the average quantity of items a worker picked per minute while working on a pick list, equals the total quantity of items in a pick list divided by how long (in minutes) it took a worker to complete the pick list.

At the end of each day, we distributed surveys to all picking workers who showed up in the experiment that day. Workers were told that their responses would be kept confidential, would not be shared with anyone else at the warehouse, and would be used exclusively for research purposes. Our daily survey collected workers' perceptions about their pick list assignment process as well as their demographics. We developed two questions to assess workers' perceived fairness about their current assignment process (relative to the alternative assignment process their peers encountered; see Table 1). (When assessing people's attitudes towards algorithmic and human-based decision-making, prior research has often had people make head-to-head comparisons of these two methods (e.g., Dietvorst et al. 2015, 2018, Longoni et al. 2019)). First, we asked workers whether they thought it would be more fair to assign pick lists using the alternative process than using their current process. Specifically, workers in the algorithm group were asked, "Do you think it would be more fair if pick lists were assigned by a human distributor?" Workers in the human group were asked, "Do you think it would be more fair if pick lists were assigned by an algorithm?" Workers in both groups responded using a five-point Likert scale from 1 ("Definitively would") to 5 ("Definitively would not"). In both groups, choosing a higher value (relative to a lower value) indicates that the worker viewed their current assignment process more favorably and less strongly believed the alternative assignment process would be more fair.

Second, we asked workers, "Which assignment process do you think is more appropriate if you were paid by item instead of by time?" (with the five-point scale ranging from 1 = "Definitively algorithmic assignment" to 5 = "Definitively human-based assignment"). We framed the question this way because people are generally sensitive to fairness in task assignment when receiving performance-based incentives (Isaac 2001). Thus, we expected that workers would report what they deemed as a fairer assignment process when they were asked to pick their preferred assignment process under a piece-rate pay scheme. For workers in both groups, choosing a higher number in response to our second question indicates that the worker viewed human-based assignment more favorably. Since we wanted to compare between groups how fair workers believed their *current assignment process* to be, we reverse coded the responses of workers in the algorithm group so that a higher value instead would indicate that the worker viewed *algorithmic* assignment—their current assignment process—as more fair. Specifically, we used six to subtract the original answer of each worker in the algorithm group. For example, if a worker in the algorithm group gave an answer of one, the worker's reverse-coded answer would be five. Reverse coding scale items is a common practice in Psychology and other fields that use survey responses (e.g., Lachman and Weaver 1998, Haimovitz and Dweck 2016). For workers in the human group, we made no adjustment to their original answers. In the end, for workers in both groups, a higher (vs. lower) value indicates that the worker viewed their current assignment process as more fair than the alternative process.

Workers' responses to these two questions (after we reverse coded the second question) are significantly and positively correlated ($r = 0.31$; $p = 0.001$). For each worker each day, we averaged her responses to these two questions to measure the extent to which she perceived her current assignment process as fairer than the alternative process (Perceived Fairness). To facilitate the interpretation of how fairness perceptions affect productivity, we constructed Standardized Perceived Fairness, which equaled Perceived Fairness divided by its standard deviation in the whole sample. Moving forward in this paper, we report results using this standardized measure.

To evaluate the convenience of their assignment process, we asked workers, "How convenient do you feel it is to receive your pick lists today?" Workers responded to this question using a five-point Likert scale (from 1 = "Very convenient" to 5 = "Very inconvenient"). We reverse coded their answers such that a higher value indicates greater convenience. As explained in Section 4.1, we were careful to ensure that it was not easier to receive pick lists in one group than in the other; otherwise, it could create an alternative explanation for productivity differences between groups. Thus, we asked workers this question to confirm no differences in this aspect.

Table 1 Measures of Fairness Perceptions

Question number	Group	Question wording
1	Algorithm	Do you think it would be more fair if pick lists were assigned by a human distributor? (1=Definitively would, 5=Definitively would not)
	Human	Do you think it would be more fair if pick lists were assigned by an algorithm? (1=Definitively would, 5=Definitively would not)
2	Algorithm	Which assignment process do you think is more appropriate if you were paid by item instead of by time? (1=Definitively algorithmic assignment, 5=Definitively human-based assignment)
	Human	Which assignment process do you think is more appropriate if you were paid by item instead of by time? (1=Definitively algorithmic assignment, 5=Definitively human-based assignment)

Note: The English translation does not match the Chinese version of our survey word by word, but it captures the meaning of our survey questions and scales well after considering the context.

To evaluate workers' emotional sensitivity to task difficulty, we asked workers how often they would feel upset if they received pick lists that were difficult to handle. Workers responded to this question using a five-point Likert scale (from 1 = "Always" to 5 = "Never"). We reverse coded their answers to this question such that a higher value indicates that the worker was more sensitive to task difficulty. We used this variable for an analyses of heterogeneous treatment effect.

To collect demographics, we asked workers to provide their gender (female or male), education (middle school or under, high school, or college or above), residence (rural or urban), and age. Since two workers did not report residence information (both of whom only worked one day during our experiment), when we add demographic controls to regressions, two observations are dropped from our regressions predicting perceived fairness and 71 observations are dropped from our regressions predicting picking efficiency.

4.3. Randomization Check

To confirm that our randomization process was successful, we compare workers' demographics and the number of days they came to work in the warehouse between the algorithm and human groups. As shown in Panel A of Table 2, the proportion of females, education levels, the proportion of workers born in urban areas, age, and the number of work days during our experiment do not significantly differ between two groups. The Kolmogorov–Smirnov test further shows that the distributions of age and work days (the only two continuous demographics variables) are comparable between the two groups of workers. These findings suggest that we have a comparable sample of workers between groups and thus our randomization process was successful.

As explained earlier, we tried to keep it equally convenient to receive pickbills in the algorithm and human groups. Indeed, workers' survey responses confirmed that workers found their pickbill

Table 2 Randomization Check

	Pick list assignment process		*Statistical test*		
	Human-based assignment	Algorithmic assignment	p-value of t-test	p-value of prop-test	p-value of ks-test
	(1)	(2)	(3)	(4)	(5)
Panel A: Worker characteristics and perceived convenience					
Gender	0.40	0.44	–	0.77	–
	(0.50)	(0.51)			
Education	1.52	1.60	0.69	–	–
	(0.65)	(0.76)			
Residence	0.25	0.13	–	0.27	–
	(0.44)	(0.34)			
Age	29.36	24.88	0.08	–	0.28
	(9.17)	(8.35)			
Number of work days	2.40	2.00	0.41	–	0.91
	(2.00)	(1.35)			
Process convenience	3.72	4.08	0.15	–	0.70
	(0.98)	(0.76)			
Observations	25(24 for residence)	25(24 for residence)	–	–	–
Panel B: Pick list characteristics					
Pick list size	20.09	21.04	0.13	–	0.12
	(20.48)	(20.62)			
Number of stocking positions	7.10	7.34	0.25	–	0.43
	(6.41)	(7.30)			
Inventory area	0.63	0.63	–	0.74	–
	(0.48)	(0.48)			
Observations	2,474	2,012	–	–	–

Note: Standard deviations are reported in the parentheses.

Note: The categorical variables in the table are defined as follows: gender: 0-male, 1-female; education: 1-middle school or under, 2-high school, 3-college or above; residence: 0-rural, 1-urban; inventory area: 0-inventory area 1, 1-inventory area 2.

Note: Process convenience: for workers that worked for more than one day during our experiment, we took the average of their responses across days to get an average measure of process convenience.

Note: We also perform randomization checks on pick list characteristics using OLS regressions. We predict pick list characteristics as a function of each worker's assignment group, following specification (2) that we describe later. All p-values are greater than 0.3.

assignment process similarly convenient between two groups (Panel A in Table 2). Moreover, as mentioned earlier, an important feature of our experiment is that pick lists were distributed to workers in two groups using the same underlying process. To confirm this was indeed the case, we compare the characteristics of pick lists received by workers in the algorithm versus human group. As our design intended, key pick list characteristics—pick list size, the number of stocking positions, and inventory area—are quite similar between groups (Panel B of Table 2), confirming that the two groups of workers actually received pick lists of the same nature.

5. Main Results from Our Main Field Experiment

5.1. The Effect of Algorithmic Assignment on Perceived Fairness

We first test whether assigning pick lists via an algorithm boosts workers' perceived fairness about their task assignment process, relative to assigning pick lists via a human (Hypothesis 1). To test this hypothesis, we apply the following regression specification to worker-day-level observations, with each observation representing worker i on day t:

$$Standardized\ perceived\ fairness_{it} = \eta_0 + \eta_1 Algorithm_i + \eta_2 X_i + \lambda_t + \epsilon_{it}, \tag{1}$$

where $Standardized\ perceived\ fairness_{it}$ refers to worker i's standardized perceived fairness on day t, $\mathrm{Algorithm}_i$ is a binary variable equaling one if worker i was in the algorithm group and zero if worker i was in the human group, and X_i is the vector of demographics controls including worker i's gender, education, residence, and age. λ_t captures day fixed effects. We cluster standard errors at the worker level. We analyze fairness at the worker-day level because this is our most granular level of observation for capturing fairness (given that each worker provided their fairness perceptions once each work day).

We report results from specification (1) and its variants (with or without controls) in Table 3. In Column 1 (without any control variables), a positive and significant coefficient on the indicator *Algorithm* (p-value < 0.0001) indicates that receiving pick lists from an algorithm significantly increases workers' perceived fairness about their assignment process, compared to receiving pick lists from a human distributor. Specifically, algorithmic assignment (relative to human-based assignment) increases perceived fairness by 0.94 standard deviations. This effect is robust and even becomes slightly larger when we control for day fixed effects (0.96 standard deviations, p-value < 0.0001; Column 2) as well as when we control for both day fixed effects and worker demographics (1.02 standard deviations, p-value < 0.0001; Column 3). Overall, these results support Hypothesis 1 that assigning pick lists by an algorithm (vs. a human) boosts workers' perceived fairness about their pick list assignment process.

We suspect that the positive effect of algorithmic assignment on fairness perceptions occurs because in our labor-intensive working environment where tasks are easier to be quantified, picking workers hold a strong equality motive for task assignments. To test this intuition, we distributed a survey to workers involved in our second field experiment (from December 27, 2019 to January 5, 2020) when they started their shift (see Online Appendix C). We asked workers whether they believed it is more important to ensure equality in task assignments or to customize task assignments based on workers' personal characteristics. Workers responded to this question using a five-point Likert scale from 1 ("Definitely prefer equality") to 5 ("Definitely prefer consideration of personal

Table 3 The Effects of Algorithmic (vs. Human-based) Assignment on Perceived Fairness and Productivity

Dependent variable	*Standardized perceived fairness*			*Picking efficiency*		
	(1)	(2)	(3)	(4)	(5)	(6)
Algorithm	0.94^{****}	0.96^{****}	1.02^{****}	0.70^{**}	0.68^{**}	0.76^{**}
	(0.20)	(0.20)	(0.23)	(0.27)	(0.23)	(0.26)
Day fixed effects	No	Yes	Yes	No	Yes	Yes
Hour fixed effects	No	No	No	No	Yes	Yes
Demographics controls	No	No	Yes	No	No	Yes
Observations	108	108	106	4,486	4,486	4,415

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$; $^{****}p<0.0001$. Average picking efficiency in the human group was 3.92.

factors"). Workers' average response was 2.49 (95% confidence interval [2.28,2.69]), which is significantly lower than 3, the mid-point of the scale ($p < 0.0001$). This suggests that workers in our field setting on average consider it more important to ensure equality than to take into account everyone's personal characteristics in task assignment.

To further understand why workers perceived algorithmic assignment more fair than human-based assignment when they care strongly about equality, we conducted structured interviews with 13 workers after both of our field experiments ended. When asked whether a pick list assignment process run by a human distributor would be fair or unfair, more than half of workers (N=7) indicated that a human-based assignment process might cause unfair outcomes. These workers mostly justified their judgment by mentioning that they believed human distributors are subject to personal biases. In addition, when asked whether they thought the pick list assignment process would be more or less fair if they could receive pick lists by scanning a bar code than if they could receive pick lists from a human distributor, most workers (N=10) believed that the process run by a machine would be more fair; and most of these workers (N=8) explained that they believed an algorithmic assignment process does not fall prey to human distributors' personal preferences, would be able to deliver equal treatments across workers, and would not selectively favor or disadvantage certain workers. We present details about our interviews in Online Appendix D.

5.2. The Effect of Algorithmic Assignment on Productivity

We next test whether assigning pick lists via an algorithm (vs. a human) enhances workers' productivity (Hypothesis 2). To test this hypothesis, we apply the following specification to pick list observations:

$$Picking\ efficiency_{ikt} = \delta_0 + \delta_1 Algorithm_i + \delta_2 X_i + \lambda_t + \epsilon_{ikt}, \tag{2}$$

where $Picking\ efficiency_{ikt}$ refers to the quantity of items worker i picked per minute for pick list k at time t, and $Algorithm_i$ and X_i are defined the same as in specification (1). In addition to day fixed effects, λ_t also includes hour fixed effects since pick list characteristics often change across hours within a day. We cluster standard error at the worker-hour level. We analyze productivity at the pick-list level because this is our most granular level of observation for capturing picking efficiency.

As shown in Columns 4-6 in Table 3, the coefficient on the indicator *Algorithm* is positive and statistically significant (all p-values < 0.01) with or without controls, which means that the algorithmic assignment treatment significantly improves workers' productivity. Specifically, without control variables, we estimate that assigning pick lists via an algorithm increases worker productivity by 17.86%, relative to the average picking efficiency of 3.92 in the human-based assignment group (Column 4). When we control for day and hour fixed effects, the effect size decreases slightly: the percentage increase in productivity caused by algorithmic assignment (relative to human-based assignment) is 17.35% (Column 5). This effect is robust and even slightly larger when we add demographics controls (19.39%; Column 6).

5.3. Average Treatment Effect of Perceived Fairness on Productivity

Next, we estimate how workers' perceived fairness about their work assignment process affects their productivity. To causally estimate this effect, we take the instrumental variable (IV) approach and use the following specifications to explain our IV estimation:

$$\textit{Picking efficiency}_{ikt} = \alpha_0 + \alpha_1 \textit{Standardized perceived fairness}_{ikt} + \alpha_2 X_i + \lambda_t + \epsilon_{ikt} \tag{3}$$

and

$$\textit{Standardized perceived fairness}_{ikt} = \beta_0 + \beta_1 \textit{Algorithm}_i + \beta_2 X_i + \lambda_t + \epsilon_{ikt}. \tag{4}$$

Directly using specification (3) to estimate the effect of fairness perceptions on productivity does not yield a causal estimate because of the omitted variable bias. Unobserved variables, such as worker ability, can be correlated with both how fair workers believe they are treated and their productivity. Therefore, we use the random assignment of workers to the algorithm group as an IV for their fairness perceptions. The two-stage least squares estimate is given in specifications (3) and (4), and standard errors are clustered at the worker-hour level. Though workers reported perceived fairness once each work day (which is why specification (1) has the notation $\textit{Standardized perceived fairness}_{it}$), here we use $\textit{Standardized perceived fairness}_{ikt}$ as the notation to indicate the level of observation (i.e., pick-list level) used in the two-stage least squares estimation.

To validate our IV estimation, we first check the *relevance assumption*: the IV Algorithm_i should be correlated with the independent variable $\text{Standardized perceived fairness}_{ikt}$. Table 3 shows that algorithmic (vs. human-based) assignment process can significantly affect workers' perceived fairness. We also confirms that this effect is statistically significant under specification (4) used in our IV estimation (all p-values < 0.0001). In addition, our IV passes the weak instrument test (F $= 1477.71$).

We next check the *exclusion restriction assumption*, which requires that the IV Algorithm$_i$ be independent of the ϵ_{ikt} in specification (3). That is, assigning pick lists by an algorithm (vs. a human) should only affect productivity by altering the workers' fairness perceptions and should not be correlated with any other factors that influence productivity. This assumption is satisfied for two reasons. First, since we randomly assigned workers to receive pick lists from either an algorithm or a human distributor, Algorithm$_i$ by design should not be correlated with variables whose value was determined prior to the experiment (e.g., worker characteristics), which we indeed verify in Section 4.3.

Second, while it is impossible to statistically prove, we carefully designed our experiment to ensure that our experimental manipulation does not affect productivity via other mechanisms than fairness perceptions. During our structured interviews, we asked picking workers, "what factors usually influence your motivation and productivity?" The most frequently mentioned factors, brought up by 7 out of 13 workers, involve pick list characteristics including the number of items they have to collect and how many stocking positions they have to get products from. As explained in Section 4.1, we worked hard to ensure that human-based assignment and algorithmic assignment essentially used the same underlying rule, and we confirmed that the key pick list characteristics that workers in our interviews highlighted—pick list size, the number of stocking positions, and inventory area—are comparable between the algorithm and human groups (Panel B of Table 2). Another factor, which was brought up by 2 out of 13 workers, is the convenience of obtaining pick lists. As explained in Section 4.1, we tried to make it equally convenient to obtain pick lists between groups. We had workers go to the same location to get their pick lists so we could keep walking distance in the assignment process the same between groups. As shown in Panel A of Table 2, workers found it similarly convenient to receive pick lists in the algorithm and human groups. Furthermore, we also asked workers, "besides pick list characteristics and the assignment process, what other factors may influence your productivity?" Factors brought up by workers include special circumstances (whether certain products are out of stocks, whether picking carts are temporarily unavailable), physical work environment (warehouse temperature, weather), and workers' physical well-being. All these factors should be comparable between two groups of workers since they worked in the same environment and were randomly assigned to the algorithm or human group.

Table 4 shows the average treatment effect of perceived fairness on productivity using IV estimation. We consistently find that workers' perceived fairness has a positive effect on productivity regardless of whether we include control variables (all p-values < 0.01 in Columns 1-3). Specifically, as perceived fairness increases by one standard deviation, worker productivity is estimated to

Table 4 IV-Estimated Effect of Perceived Fairness on Productivity

Dependent variable	*Picking efficiency*		
	(1)	(2)	(3)
Standardized perceived fairness	0.72**	0.68**	0.68**
	(0.27)	(0.23)	(0.24)
Day fixed effects	No	Yes	Yes
Hour fixed effects	No	Yes	Yes
Demographics controls	No	No	Yes
Observations	4,486	4,486	4,415

Note: *p<0.05; **p<0.01; ***p<0.001. Average picking efficiency across algorithm and human groups was 4.24.

significantly increase by 16.04%-16.98%, relative to the average pick efficiency of 4.24 across both algorithm and human groups (Columns 1-3).

6. Heterogeneous Treatment Effects from Our Main Field Experiment

In this section, we test how the effects of algorithmic (vs. human-based) assignment on fairness perceptions and productivity vary across workers. Understanding which type of worker experiences greater productivity gains under algorithmic assignment may not only shed further light on why algorithmic assignment boosts productivity in our setting but also suggest to managers the segment of workers to whom they could first apply algorithmic assignment processes.

6.1. Sensitivity to Task Difficulty

We first test whether workers with a higher sensitivity to task difficulty may exhibit a larger boost in productivity when they receive pick lists from an algorithm than when they receive pick lists from a human distributor. We suspect that workers who tend to feel upset about receiving difficult tasks care more about whether they receive tasks from a fair process and are more likely to adjust their behavior based on how fairly they think they are treated, compared to workers who tend not to feel upset about getting difficult tasks. Therefore, fairness perceptions about a task assignment process should have a larger positive impact on productivity among workers who are more (vs. less) upset about getting difficult tasks. Combining this argument with our Hypothesis 1 that workers in our setting perceive an algorithmic assignment process as more fair than a human-based assignment process, we suspect that workers with a high sensitivity to task difficulty would show a larger productivity difference between algorithmic and human-based assignment processes, compared to workers with a low sensitivity to task difficulty.

To test this prediction, we construct a measure of sensitivity to task difficulty using workers' responses to our survey question that asked them how often they felt upset when they received difficult tasks. (One worker didn't answer this question and was thus dropped from our analysis involving sensitivity to task difficulty. This worker was also one of the two workers who did not

Table 5 HTE Based on Sensitivity to Task Difficulty

	Subsamples of workers:					
	High sensitivity			Low sensitivity		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: The effect of algorithmic assignment on perceived fairness						
Dependent variable	*Standardized perceived fairness*					
Algorithm	1.26****	1.24****	1.60****	0.78*	0.77	1.15*
	(0.28)	(0.28)	(0.32)	(0.32)	(0.44)	(0.45)
Day fixed effect	No	Yes	Yes	No	Yes	Yes
Demographics controls	No	No	Yes	No	No	Yes
Observations	54	54	53	53	53	53
Panel B: The effect of algorithmic assignment on productivity						
Dependent variable	*Picking efficiency*					
Algorithm	1.19***	1.06***	1.48****	0.10	0.19	−0.24
	(0.39)	(0.34)	(0.37)	(0.34)	(0.32)	(0.30)
Panel C: Average treatment effect of perceived fairness on productivity						
Dependent variable	*Picking efficiency*					
Standardized perceived fairness	1.04***	0.93***	0.97****	0.11	0.20	−0.18
	(0.35)	(0.30)	(0.23)	(0.38)	(0.33)	(0.23)
Day fixed effect	No	Yes	Yes	No	Yes	Yes
Hour fixed effect	No	Yes	Yes	No	Yes	Yes
Demographics controls	No	No	Yes	No	No	Yes
Observations	2,361	2,361	2,311	2,104	2,104	2,104

Note: *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001. Average picking efficiency in the human group was 3.90 for workers with a higher sensitivity to task difficulty and 3.95 for workers with a lower sensitivity; average picking efficiency across algorithm and human groups was 4.49 for workers with a higher sensitivity to task difficulty and 3.99 for workers with a lower sensitivity.

provide residence information.) For workers that worked for more than one day, we took the average of their responses across days. The median of this measure was 2 across workers. We split our sample based on the median such that workers whose average response was higher than 2 were in the "high sensitivity" category (N=21) and workers whose average response was equal or lower than 2 were in the "low sensitivity" category (N=29). For these two subsamples of workers, we separately estimate the effect of algorithmic (vs. human-based) assignment on perceived fairness using specification (1) and on productivity using specification (2), as well as the average treatment effect of perceived fairness on productivity using specifications (3)-(4). We report the results of these regressions in Table 5.

As shown in Panel A of Table 5, assigning pick lists via an algorithm (vs. a human distributor) significantly increases workers' fairness perceptions about their current assignment process by 1.24-1.60 standard deviations (depending on the inclusion of control variables) among workers with a high sensitivity to task difficulty (all p-values < 0.0001 in Columns 1-3). The effect is directionally weaker but still generally holds among workers with a low sensitivity to task difficulty: among this subsample, algorithmic assignment increases perceived fairness by 0.78-1.15 standard deviations, relative to human-based assignment (p-values < 0.05 in Column 4 without controls and Column 6 with full controls).

As shown in Panel B of Table 5, algorithmic assignment boosts the productivity of workers with a high sensitivity to task difficulty by 27.18%-37.95%, relative to the average picking efficiency of 3.90 among workers with a high sensitivity in the human group (all p-values < 0.001 in Columns 1-3). However, algorithmic (vs. human-based) assignment does not significantly impact productivity of workers with a low sensitivity to task difficulty (Columns 4-6).

Furthermore, Panel C of Table 5 indicates that fairness perceptions have a stronger impact on workers with a high sensitivity to task difficulty than on workers with a low sensitivity. Specifically, for workers with a high sensitivity, increasing perceived fairness by one standard deviation can lead to an increase in productivity by 20.71%-23.16%, relative to the average pick efficiency of 4.49 across all high-sensitivity workers (all p-values < 0.001 in Columns 1-3). However, the effect of fairness on productivity is close to zero in magnitude and not statistically significant among those with a low sensitivity (Columns 4-6). Altogether, we find that workers with a higher (vs. lower) sensitivity to task difficulty exhibit a bigger productivity lift when they receive pick lists from an algorithm than from a human distributor. This increase in productivity is likely to be due to the bigger role that fairness perceptions play in driving more sensitive workers' productivity.

6.2. Workers' Education Level

We next explore whether workers' education levels affect how they respond to algorithmic assignment. People with higher levels of education tend to have higher self-esteem, maintaining a more positive evaluation of their own worth, value, and importance, compared with these with lower levels of education (Twenge and Campbell 2002). Further, people with high self-esteem, relative to these with low self-esteem, are more eager to embrace fair treatments and more likely to adjust their attitudes and effort at work based on their fairness perceptions (Wiesenfeld et al. 2007). Connecting these arguments, we speculate that fairness perceptions about task assignment process would have a larger impact on productivity for workers with a higher education level than for workers with a lower education level. Thus, workers with a higher (vs. lower) level of education may improve productivity by a larger degree when they receive pick lists from an algorithm than from a human.

We split our sample by education level: workers whose education level is at or below middle school form a subsample (N=28), and workers whose education is at or above high school form another subsample (N=22). We separate the sample based on whether a worker achieved a degree higher than middle school because China has the nine-year compulsory education policy: citizens

Table 6 HTE Based on Workers' Education Level

	Subsamples of workers:					
	High school or above			Middle school or under		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: The effect of algorithmic assignment on perceived fairness						
Dependent variable	*Standardized perceived fairness*					
Algorithm	1.17^{***}	1.26^{***}	1.39^{**}	0.73^{**}	0.77^{**}	0.87^{*}
	(0.33)	(0.31)	(0.40)	(0.23)	(0.27)	(0.37)
Day fixed effect	No	Yes	Yes	No	Yes	Yes
Demographics controls	No	No	Yes	No	No	Yes
Observations	52	52	52	56	56	54
Panel B: The effect of algorithmic assignment on productivity						
Dependent variable	*Picking efficiency*					
Algorithm	1.30^{***}	0.82^{**}	1.57^{****}	0.25	0.40	0.32
	(0.40)	(0.33)	(0.35)	(0.36)	(0.31)	(0.40)
Panel C: Average treatment effect of perceived fairness on productivity						
Dependent variable	*Picking efficiency*					
Standardized perceived fairness	1.08^{***}	0.66^{**}	1.09^{****}	0.30	0.49	0.36
	(0.35)	(0.27)	(0.28)	(0.43)	(0.38)	(0.44)
Day fixed effect	No	Yes	Yes	No	Yes	Yes
Hour fixed effect	No	Yes	Yes	No	Yes	Yes
Demographics controls	No	No	Yes	No	No	Yes
Observations	2,177	2,177	2,177	2,309	2,309	2,238

Note: *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001. Average picking efficiency in the human group was 3.84 for workers whose education level was at or above high school and 4.03 for workers whose education level was at or below middle school; average picking efficiency across algorithm and human groups was 4.33 for workers whose education level was at or above high school and 4.15 for workers whose education level was at or below middle school.

are required to attend school for at least nine years including six years of primary education and three years of middle school.

As shown in Panel A of Table 6, for both subsamples of workers, workers under algorithmic assignment found their assignment process more fair than workers under human-based assignment. Specifically, assigning pick lists via an algorithm (vs. via a human distributor) significantly increases workers' fairness perceptions of their current assignment process by 1.17-1.39 standard deviations among workers with a high school degree or above (all p-values < 0.005 in Columns 1-3) and by 0.73-0.87 standard deviations among workers without a high school degree (all p-values < 0.05 in Columns 4-6).

As shown in Panel B of Table 6, for workers with a high school degree or above, algorithmic assignment boosts productivity by 21.35%-40.89%, relative to the average picking efficiency of 3.84 among workers with a high school degree or above in the human group (all p-values < 0.01 in

Columns 1-3). However, algorithmic (vs. human-based) assignment has no significant effect on the productivity of workers without a high school degree (Columns 4-6).

Furthermore, Panel C of Table 6 indicates that fairness perceptions have a stronger impact on workers with a high school degree or above. Specifically, for these workers, increasing perceived fairness by one standard deviation can increase productivity by 15.24%-25.17% relative to the average picking efficiency of 4.33 across all workers with a high school degree or above (all p-values < 0.01 in Columns 1-3). However, the effect of fairness perceptions on productivity is close to zero in magnitude and not statistically significant among those without a high school degree. Altogether, we find that workers with a higher (vs. lower) education level exhibit a bigger productivity lift when they receive pick list assignments from an algorithm than from a human distributor. This increase in productivity can be attributed to the greater impact of fairness perceptions on more educated workers' productivity.

7. Online Experiments Assessing the Impact of Algorithmic Assignment on Fairness Perceptions

Following our field experiments, we conducted two additional online scenario-based experiments with survey respondents in the United States to replicate the effect of algorithmic (vs. human-based) assignment on perceived fairness. We designed online experiments to complement our field experiments in a few ways. First, we intended to demonstrate the generalizability of our findings about fairness perceptions using a larger sample of participants under a different culture than our field experiments. Second, both of our field experiments measured workers' fairness perceptions about their current assignment process by asking them to compare the algorithmic assignment process with the human-based assignment process. In our online experiments, we measured people's perceptions about their current assignment process without drawing any comparison with alternative assignment processes. Third, although we adopted the best design possible in our field experiments, workers could communicate with each other across groups; in the online experiments, participants did not know about alternative assignment processes. Our two online experiments followed the same design with one exception and yielded consistent results. We report one of the experiments below and detail the other experiment in Online Appendix E. Online experiments have often been used to complement field studies (Derfler-Rozin et al. 2016, Buell et al. 2017, Staats et al. 2018). In their recent book chapter about field experiments, Ibanez and Staats (2018) highlighted that "lab and field should be seen as complements rather than substitutes; in particular... researchers can go back to the lab after field experiments" (p. 125).

7.1. Experimental Design and Analysis

We recruited study participants from an online labor marketplace, Amazon's Mechanical Turk, to complete a 4-minute study in exchange for \$0.60. Only people who accessed our study on a non-mobile device, successfully completed a CAPTCHA, and passed an attention check were allowed to start our study. People who satisfied these criteria were asked to imagine themselves as a warehouse picking worker and read about descriptions of picking tasks. To ensure that participants understood this work setting and could immerse themselves into the scenario, we required that participants had to correctly answer three questions about the scenario in order to continue with the study. Those who passed our comprehension check questions and completed our study (N=201; 41.29% female, $M_{age} = 39.47$) comprised our final sample.

Upon passing our comprehension check questions, participants were randomly assigned to either the algorithm condition (N=100) or the human condition (N=101). The descriptions of the two conditions mimicked the set-up in our field experiments. In the algorithm condition, participants were told that their pick list assignment process was run by a machine and that they received pick lists by scanning a bar code marked at the distribution station. In the human condition, participants were told that their pick list assignment process was run by a human and that they received hard-copy pick lists from a manager at the distribution station. Then participants in both conditions were told that the average pick list size (or the average number of items in a pick list) in the warehouse was 21 (based on the actual average pick list size in our main field experiment). Participants were also presented with their average pick list size on each of the past 10 workdays, and this information was the same between the algorithm and human conditions. We presented this information about pick list size and kept it the same between conditions so as to control for pick list assignment outcomes and cleanly investigate people's perceptions of a given assignment process, as we did in the field. In our other online experiment, we did not provide information about pick list size and obtained similar results (see Online Appendix E).

One assumption underlying our Hypothesis 1 is that people believe humans are subject to personal biases and algorithmic assignment processes are more capable of delivering equal treatments across workers. In our field setting, most workers in our interviews did express such beliefs (as we discussed in Section 5.1). To test this assumption in our online experiment, we asked participants to indicate their agreement with the following statement about the assignment process they imagined getting pick lists from (either algorithmic or human-based): "I think this assignment process would treat every worker perfectly equally" (from 1 = "Strongly disagree" to 7 = "Strongly agree"). Choosing a higher (vs. lower) value indicates that the participant viewed their assignment process as more capable of preserving equality.

Then we assessed fairness perceptions about an assignment process by asking participants to indicate their agreement with four statements adapted from Conlon et al. (2004) and Newman et al. (2020): (1) "the way this warehouse assigns pick lists seems fair," (2) "the warehouse's process for distributing pick lists is fair," (3) "the decision regarding whether I get more difficult pick lists is fair," and (4) "the outcome of the pick list distribution is fair." The anchors on the scale ranged from 1 ("Strongly disagree") to 7 ("Strongly agree"). Choosing a higher (vs. lower) value indicates that the participant viewed their assignment process as more fair. Participants' ratings of these four statements reached a high inter-item reliability (Cronbach's alpha = 0.96) and were thus averaged to form a composite score of *Perceived Fairness*. Following the analysis in our field experiment, we constructed *Standardized Perceived Fairness*, which equaled *Perceived Fairness* divided by its standard deviation in the whole sample.

Next, to check whether people in our online experiment held a strong equality motive for the assignment of picking tasks, we measured the perceived importance of equality and uniqueness. Specifically, we asked participants to separately rate how important they thought it was for a pick list assignment process to treat all workers equally and how important it was for a pick list assignment process to take into account individual workers' characteristics. The anchors on both scales ranged from 1 ("Not important at all") to 7 ("Very important"). Choosing a higher value indicates a higher perceived importance. In addition, we also used a single item as in our field setting and asked participants which of the two objectives they thought the warehouse should prioritize when it comes to assign picking tasks: treating all workers equality or taking into consideration personal characteristics. We obtained consistent results using these two methods to examine the relative importance of the equality versus uniqueness motive (see Online Appendix E), and focus on the former in the paper. Finally, participants reported their gender, age, and education.

7.2. Results

By comparing participants' importance ratings for equality versus uniqueness, we first confirm that people on average prioritize equality over uniqueness in the warehouse task assignment setting ($M_{equality} = 5.88$, $SD = 1.27$ vs. $M_{uniqueness} = 4.53$, $SD = 1.71$; t(200.00) = 8.84, p-value < 0.0001 for a paired t test, Cohen's d = 0.90). Second, supporting the assumption underlying our Hypothesis 1, people view the assignment process run by a machine as more capable of preserving equality than the assignment process run by a human ($M_{algorithm} = 4.95$, $SD = 1.12$ vs. $M_{human} = 4.36$, $SD = 1.13$; t(198.61) = 2.74, p-value < 0.001, Cohen's d=0.73). Further, in support of Hypothesis 1, participants in the algorithm condition perceived their assignment process more fair than those in the human condition ($M_{algorithm} = 4.23$, $SD = 0.99$ vs. $M_{human} = 3.79$, $SD = 0.96$; t(198.70) = 3.20, p-value < 0.005, Cohen's d=0.45). Since the variances are unequal between groups, we report in the paper degrees of freedom that have been adjusted for variance.

8. General Discussion

We study the impact of algorithmic (vs. human-based) work assignment on assignment recipients' fairness perceptions and productivity. In two randomized field experiments, we randomly assigned picking workers in one of Alibaba's warehouses to receive tasks either from an algorithm or from a human distributor. Combining performance data from Alibaba's digital labor system with survey responses, we present several key findings.

First, we find that assignment recipients' fairness perceptions change with the framing of how their tasks are determined. In our field setting where workers believe that task assignments should prioritize equality over consideration of personal characteristics, receiving tasks from an algorithm increases workers' perceived fairness by 0.94-1.02 standard deviations (depending on the inclusion of control variables), relative to receiving tasks of an identical nature from a human. While we sought to ensure that tasks were distributed to workers in both groups using the same underlying rule, workers may believe that algorithms can apply rules consistently across workers and treat everyone equally but human distributors have the discretion to favor some workers, as our interviews suggest.

Second, we find that the two types of task assignment methods have an economically meaningful difference in productivity: workers' picking efficiency increases by 17.35%-19.39% when pick lists are assigned by an algorithm than when pick lists are assigned by a human distributor. This is driven by the positive effect of fairness perceptions on productivity. Using the IV approach, we estimate that a one-standard-deviation increase in perceived fairness boosts workers' picking efficiency by 16.04%-16.98%.

In addition, our analysis of heterogeneous treatment effects suggests that the effects of algorithmic (vs. human-based) assignment on productivity are more prominent among workers who reported being more (vs. less) upset about receiving difficult tasks and among workers who had a higher (vs. lower) education level; these patterns may occur because the fairness of an assignment process matters more and has a stronger relationship with productivity for workers with greater task difficulty sensitivity as well as for workers with a higher education level.

We conducted two auxiliary online experiments. Similar to Chinese picking workers in our collaborating warehouse, U.S. survey respondents also believe that it is more important for a task assignment process to maintain equality than to consider task recipients' unique characteristics. They expect algorithms to be better at delivering equal treatments across workers than human task distributors. As a generalization of our finding about perceived fairness in the field experiments, our online experiments reveal that people in Western culture also perceive an algorithmic assignment process to be fairer than a human-based assignment process, even when assignment outcomes are the same between these two processes.

Our research has important practical implications. First, our results highlight that when algorithms are applied to solve operational problems, they can have broader impacts beyond offering greater efficiency and accuracy than humans. Managers may want to look for opportunities to introduce algorithmic assignment processes to reap their benefits on fairness perceptions and productivity. Second, our ability to observe productivity differences between groups even when the algorithm and human distributors assigned objectively comparable pick lists suggests that the framing and laypeople's beliefs about an assignment process matter to productivity, not just how the process actually works and what decisions it actually makes. This insight encourages managers to find the most motivating framing of task assignment processes for their workers and to understand their workers' beliefs about different allocation processes. Third, our results underscore the important role of psychological factors such as fairness perceptions in driving workers' motivation and productivity. Simple strategies like drawing employees' attention to algorithmic task assignment processes that are used in their organizations may lead employees to perceive their managers as caring about fairness, which could be beneficial for employees' performance. In addition, our heterogeneous treatment effects suggest that it is particularly useful for managers to consider applying or highlighting algorithmic decision processes if their workers are better educated.

Our research opens up avenues for future research. First, in our field experiments, we tracked workers' performance at the pick list level, only knowing the quantity of total items picked, the corresponding stocking shelves and inventory area, as well as the starting and ending times of each pick list. With the development of wearable monitors in e-commerce warehouses, it would be possible to track workers' movement and actions more precisely, such as measuring picking workers' effort by their walking speed and heart rate and assessing task difficulty by workers' movements (e.g., raising their arm, bending over). Such granular data would allow researchers and managers to more comprehensively understand how different assignment methods affect workers' behavior and physiological reactions for different types of tasks and then design interventions accordingly.

Second, our research focuses on work settings where people tend to believe that task allocations should prioritize equality. Connecting our findings with recent work suggesting that people worry algorithms cannot incorporate their personal characteristics and thus view algorithmic decision processes as less fair (Newman et al. 2020), we think it is worth studying whether the motive people prioritize, equality or uniqueness, shapes their reactions to algorithmic (vs. human) assignment processes. Our online experiments provide initial evidence for the role of motive. Specifically, we find that the positive effect of algorithmic (vs. human-based) assignment process on fairness perception holds strongly among people who prioritize equality over uniqueness but is weak among people who prioritize uniqueness over equality. See detailed results in Online Appendix E. Future work can more systematically examine how fundamental human motives like equality and uniqueness affect reactions to algorithmic decision processes.

Another fruitful future research direction is to examine how our findings apply to other settings (e.g., service, creative work) and to understand when, to whom, and in what type of work, equality (vs. uniqueness) motive matters more to workers. This will allow researchers and managers to better predict human reactions to algorithmic decision processes and identify the type of work situation where algorithmic decision-making processes can yield the largest benefits for perceived fairness and productivity.

References

Aksin, Z, Sarang Deo, Jónas Jónasson, Kamalini Ramdas. 2015. Learning from many: Partner diversity and team familiarity in fluid teams .

Angwin, Julia, Jeff Larson, Surya Mattu, Lauren Kirchner. 2016. Machine bias. *ProPublica, May* **23** 2016.

Aronow, Peter M. 2012. A general method for detecting interference between units in randomized experiments. *Sociological Methods & Research* **41**(1) 3–16.

Ball, Ryan T, Eric Ghysels. 2017. Automated earnings forecasts: Beat analysts or combine and conquer? *Management Science* **64**(10) 4936–4952.

Banerjee, Abhijit V, Shawn Cole, Esther Duflo, Leigh Linden. 2007. Remedying education: Evidence from two randomized experiments in india. *The Quarterly Journal of Economics* **122**(3) 1235–1264.

Bhandari, Atul, Alan Scheller-Wolf, Mor Harchol-Balter. 2008. An exact and efficient algorithm for the constrained dynamic operator staffing problem for call centers. *Management Science* **54**(2) 339–353.

Bloom, Nicholas, James Liang, John Roberts, Zhichun Jenny Ying. 2015. Does working from home work? evidence from a chinese experiment. *The Quarterly Journal of Economics* **130**(1) 165–218.

Buell, Ryan W, Tami Kim, Chia-Jung Tsay. 2017. Creating reciprocal value through operational transparency. *Management Science* **63**(6) 1673–1695.

Buell, Ryan W, Michael I Norton. 2011. The labor illusion: How operational transparency increases perceived value. *Management Science* **57**(9) 1564–1579.

Caliskan, Aylin, Joanna J Bryson, Arvind Narayanan. 2017. Semantics derived automatically from language corpora contain human-like biases. *Science* **356**(6334) 183–186.

Castelo, Noah, Maarten W Bos, Donald R Lehmann. 2019. Task-dependent algorithm aversion. *Journal of Marketing Research* **56**(5) 809–825.

Choudhary, Vivek, Masha Shunko, Serguei Netessine. 2020. Does immediate feedback make you not try as hard? a study on automotive telematics. *Manufacturing & Service Operations Management* .

Chouldechova, Alexandra. 2017. Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data* **5**(2) 153–163.

Cohen-Charash, Yochi, Paul E Spector. 2001. The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes* **86**(2) 278–321.

Colquitt, Jason A, Donald E Conlon, Michael J Wesson, Christopher OLH Porter, K Yee Ng. 2001. Justice at the millennium: a meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology* **86**(3) 425.

Conlon, Donald E, Christopher OLH Porter, Judi McLean Parks. 2004. The fairness of decision rules. *Journal of Management* **30**(3) 329–349.

Corbett-Davies, Sam, Emma Pierson, Avi Feller, Sharad Goel, Aziz Huq. 2017. Algorithmic decision making and the cost of fairness. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 797–806.

Cowgill, Bo. 2018. Bias and productivity in humans and algorithms: Theory and evidence from resume screening. *Columbia Business School, Columbia University* .

Cowgill, Bo, Catherine E Tucker. 2019. Economics, fairness and algorithmic bias .

Datta, Amit, Michael Carl Tschantz, Anupam Datta. 2015. Automated experiments on ad privacy settings. *Proceedings on Privacy Enhancing Technologies* **2015**(1) 92–112.

Dawes, Christopher T, James H Fowler, Tim Johnson, Richard McElreath, Oleg Smirnov. 2007. Egalitarian motives in humans. *Nature* **446**(7137) 794–796.

Derfler-Rozin, Rellie, Celia Moore, Bradley R Staats. 2016. Reducing organizational rule breaking through task variety: How task design supports deliberative thinking. *Organization Science* **27**(6) 1361–1379.

Dietvorst, Berkeley J, Soaham Bharti. 2020. People reject algorithms in uncertain decision domains because they have diminishing sensitivity to forecasting error. *Psychological Science* **31**(10) 1302–1314.

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

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

Dijkstra, Jaap J, Wim BG Liebrand, Ellen Timminga. 1998. Persuasiveness of expert systems. *Behaviour & Information Technology* **17**(3) 155–163.

Duflo, Esther, Emmanuel Saez. 2003. The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *The Quarterly Journal of Economics* **118**(3) 815–842.

Dwork, Cynthia, Moritz Hardt, Toniann Pitassi, Omer Reingold, Richard Zemel. 2012. Fairness through awareness. *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*. ACM, 214–226.

Fehr, Ernst, Lorenz Goette, Christian Zehnder. 2009. A behavioral account of the labor market: The role of fairness concerns. *Annual Reviews of Economics* **1**(1) 355–384.

Fryer Jr, Roland G, Steven D Levitt, John List, Sally Sadoff. 2012. Enhancing the efficacy of teacher incentives through loss aversion: A field experiment. Tech. rep., National Bureau of Economic Research.

Greenberg, Jerald, Jason A Colquitt. 2013. *Handbook of organizational justice*. Psychology Press.

Haimovitz, Kyla, Carol S Dweck. 2016. What predicts children's fixed and growth intelligence mind-sets? not their parents' views of intelligence but their parents' views of failure. *Psychological Science* **27**(6) 859–869.

Hossain, Tanjim, John A List. 2012. The behavioralist visits the factory: Increasing productivity using simple framing manipulations. *Management Science* **58**(12) 2151–2167.

Huckman, Robert S, Bradley R Staats, David M Upton. 2009. Team familiarity, role experience, and performance: Evidence from indian software services. *Management science* **55**(1) 85–100.

Ibanez, Maria R, Bradley R Staats. 2018. Behavioral empirics and field experiments. *The Handbook of Behavioral Operations* 121–147.

Ibanez, Maria R, Michael W Toffel. 2020. How scheduling can bias quality assessment: Evidence from food-safety inspections. *Management Science* **66**(6) 2396–2416.

Isaac, Joe E. 2001. Performance related pay: The importance of fairness. *Journal of Industrial Relations* **43**(2) 111–123.

Jago, Arthur S. 2019. Algorithms and authenticity. *Academy of Management Discoveries* **5**(1) 38–56.

Karacaoglu, Nil, Antonio Moreno, Can Ozkan. 2018. Strategically giving service: The effect of real-time information on service efficiency. *Available at SSRN 3260035* .

Kc, Diwas Singh. 2020. Heuristic thinking in patient care. *Management Science* **66**(6) 2545–2563.

Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, Sendhil Mullainathan. 2017. Human decisions and machine predictions. *The Quarterly Journal of Economics* **133**(1) 237–293.

Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan, Ashesh Rambachan. 2018. Algorithmic fairness. *AEA Papers and Proceedings*, vol. 108. 22–27.

Kleinberg, Jon, Sendhil Mullainathan, Manish Raghavan. 2016. Inherent trade-offs in the fair determination of risk scores. *arXiv preprint arXiv:1609.05807* .

Kropp, Dean H, Robert C Carlson. 1984. A lot-sizing algorithm for reducing nervousness in mrp systems. *Management Science* **30**(2) 240–244.

Lachman, Margie E, Suzanne L Weaver. 1998. The sense of control as a moderator of social class differences in health and well-being. *Journal of Personality and Social Psychology* **74**(3) 763.

Lambrecht, Anja, Catherine Tucker. 2019. Algorithmic bias? an empirical study of apparent gender-based discrimination in the display of stem career ads. *Management Science* **65**(7) 2966–2981.

Leung, Eugina, Gabriele Paolacci, Stefano Puntoni. 2018. Man versus machine: Resisting automation in identity-based consumer behavior. *Journal of Marketing Research* **55**(6) 818–831.

Leventhal, Gerald S, Jurgis Karuza, William R Fry. 1980. Beyond fairness: A theory of allocation preferences. *Justice and Social Interaction* **3**(1) 167–218.

Li, Yao, Lauren Xiaoyuan Lu, Susan F Lu, Jian Chen. 2020. The value of health it interoperability: Evidence from interhospital transfer of heart attack patients. *Available at SSRN 3557010* .

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

Longoni, Chiara, Andrea Bonezzi, Carey K Morewedge. 2019. Resistance to medical artificial intelligence. *Journal of Consumer Research* **46**(4) 629–650.

Luo, Xueming, Siliang Tong, Zheng Fang, Zhe Qu. 2019. Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science* **38**(6) 937–947.

McAfee, Andrew, Erik Brynjolfsson. 2017. *Machine, platform, crowd: Harnessing our digital future*. WW Norton & Company.

Mookerjee, Radha, Subodha Kumar, Vijay S Mookerjee. 2017. Optimizing performance-based internet advertisement campaigns. *Operations Research* **65**(1) 38–54.

Newman, David T, Nathanael J Fast, Derek J Harmon. 2020. When eliminating bias isn't fair: Algorithmic reductionism and procedural justice in human resource decisions. *Organizational Behavior and Human Decision Processes* **160** 149–167.

Obermeyer, Ziad, Brian Powers, Christine Vogeli, Sendhil Mullainathan. 2019. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* **366**(6464) 447–453.

Rai, Tage Shakti, Alan Page Fiske. 2011. Moral psychology is relationship regulation: moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review* **118**(1) 57.

Staats, Bradley R, Diwas S Kc, Francesca Gino. 2018. Maintaining beliefs in the face of negative news: The moderating role of experience. *Management Science* **64**(2) 804–824.

Sun, Jiankun, Dennis Zhang, Haoyuan Hu, Jan A Van Mieghem. 2020. Predicting human discretion to adjust algorithmic prescription: A large-scale field experiment in warehouse operations. *Management Science* .

Tan, Tom Fangyun, Serguei Netessine. 2019. When you work with a superman, will you also fly? an empirical study of the impact of coworkers on performance. *Management Science* **65**(8) 3495–3517.

Twenge, Jean M, W Keith Campbell. 2002. Self-esteem and socioeconomic status: A meta-analytic review. *Personality and Social Psychology Review* **6**(1) 59–71.

Tyler, Tom R. 1989. The psychology of procedural justice: A test of the group-value model. *Journal of Personality and Social psychology* **57**(5) 830.

Van Donselaar, Karel H, Vishal Gaur, Tom Van Woensel, Rob ACM Broekmeulen, Jan C Fransoo. 2010. Ordering behavior in retail stores and implications for automated replenishment. *Management Science* **56**(5) 766–784.

Wiesenfeld, Batia M, William B Swann Jr, Joel Brockner, Caroline A Bartel. 2007. Is more fairness always preferred? self-esteem moderates reactions to procedural justice. *Academy of Management Journal* **50**(5) 1235–1253.

Wu, Anqi Angie, Yixin Iris Wang. 2019. The more monitoring, the better quality? empirical evidence from the generic drug industry .

Xu, Chunchen, Arthur S Jago. 2020. Algorithmic decision making undermines affective commitment. *Academy of Management Proceedings*, vol. 2020. Academy of Management Briarcliff Manor, NY 10510, 15149.

Xu, Yuqian, Lingjiong Zhu. 2020. Operational risk management: Team-based effort and incentive bonus. *Available at SSRN 3191887* .

Zhang, Yu, Vidyadhar Kulkarni. 2018. Automated teller machine replenishment policies with submodular costs. *Manufacturing & Service Operations Management* **20**(3) 517–530.

Online Appendix

Online Appendix A: Pick List Example

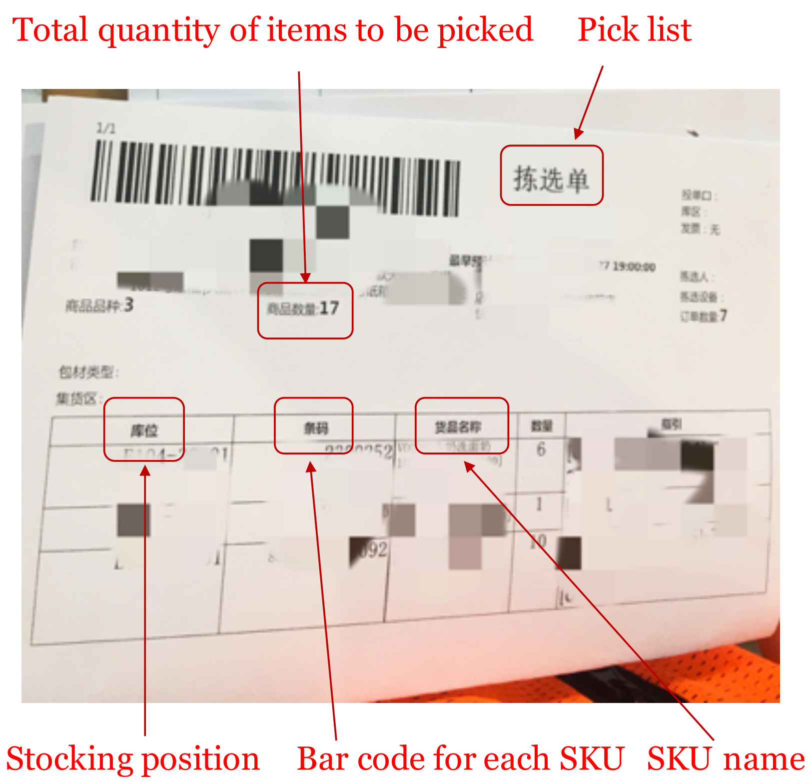


Figure 3 **Pick List Example**

Online Appendix B: Check Interference Between Workers via a Statistical Test

Following Aronow (2012), we use a statistical test to check whether the behavior of a particular worker depends only on her assignment process, not on the assignment process of others working around her. This is an ex post method to detect interference between units (where each unit represents a worker in our context) in a randomized experiment. In our setting, interference between workers could occur if (1) workers in the human group may view their assignment process as less fair and become less motivated (relative to workers in the algorithm group) because they learned that workers in the other group scanned a bar code to get their pick lists; or (2) if workers in the algorithm group may perceive their assignment process more fair and become more motivated (relative to workers in the human group) because they learned that workers in the other group got their pick lists from a human distributor.

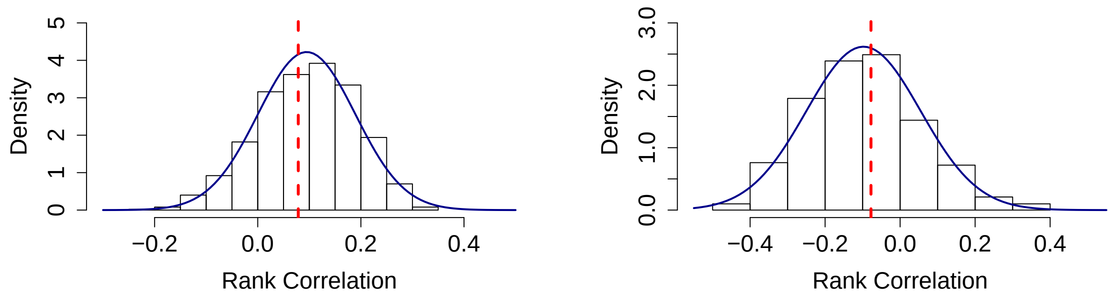
To perform the test recommended by Aronow (2012), we first randomly select 12 workers from the human condition as the fixed subset. The remaining 38 workers belong to our variant subset.

Then we draw 1,000 simulations on the experimental condition of the variant subset. In each simulation, we randomly select 25 workers from the variant subset to be in the algorithm condition and 13 workers to be in the human condition. For each simulation and for each day during our field experiment, we calculate the simulated daily algorithmic treatment rate, which equals the proportion of workers who were assigned to the algorithm condition *in the simulation* among all workers coming to work that day. Then for each simulation, we compute the Spearman's rank correlation coefficient ρ between the productivity of each worker in the fixed subset and the simulated daily algorithmic treatment rate that day.

Across 1,000 simulations, we obtain 1,000 values of ρ. We plot the distribution of ρ in Figure 4(a). Since workers in the variant subset were purely randomly assigned to the algorithm vs. human condition in each simulation and 1,000 simulations were independent, Figure 4(a) presents the approximate distribution of ρ associated with the null hypothesis that interference on productivity between units did not occur for workers in the human group. The dashed line in Figure 4(a) represents the observed correlation coefficient ρ in our first field experiment. The observed ρ is around the center of null distribution, yielding $p = 0.41$. (The p-value equals the min of two areas: the area to the left of the dashed line and the area to the right, since this is a two-sided test and we do not know a priori whether the observed correlation would be negative or positive.) Therefore, we can't reject the null hypothesis that there is no interference on productivity for workers in the human group.

Using these 1,000 simulations, we perform a similar test on perceived fairness. For each simulation, we compute the Spearman's rank correlation coefficient ρ between the perceived fairness of each worker in the fixed subset on a day and the simulated daily algorithmic treatment rate that day. We plot the distribution of ρ in Figure 4(b). The observed ρ, as indicated by the dashed line, is again close to the center of sharp null distribution, yielding $p = 0.43$ (i.e., the area to the right of the dashed line, which is smaller than the area to the left). Thus, we can't reject the null hypothesis that there is no interference on fairness perceptions for workers in the human group.

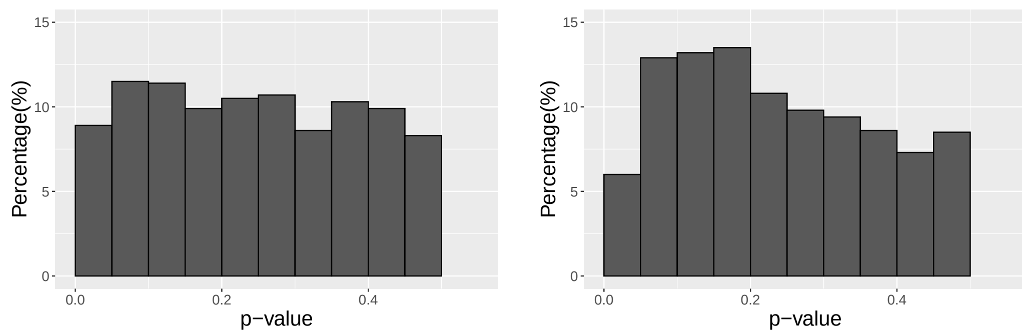
So far, we have done 1,000 simulations by taking the same set of 12 workers in the human group as the fixed subset. To confirm the robustness of our test, we randomly draw 12 workers from 25 workers in the human group as the fixed subset for 1,000 times. Each time we randomly select a fixed subset, we repeat the process described above involving 1,000 simulations and obtain two p-values (one for productivity and one for perceived fairness). Figure 5(a) shows the distribution of p-values for the interference test on productivity across 1,000 draws of fixed subsets. Note that as explained above, based on how Aronow (2012) calculate p-values, p-values are between 0 and 0.5. The p-values from our 1,000 draws are smaller than 0.05 only 8.60% of the time, lower than 10% (the chance level for p-values to fall below 0.05 under uniform distribution since p-values are



(a) Algorithmic Treatment Rate and Productivity (Actually Observed $\rho = 0.08$)

(b) Algorithmic Treatment Rate and Perceived Fairness (Actually Observed $\rho = -0.08$)

Figure 4 Distribution of Rank Correlation Coefficients Across 1,000 Simulations and the Observed Coefficient



(a) Interference Test on Productivity

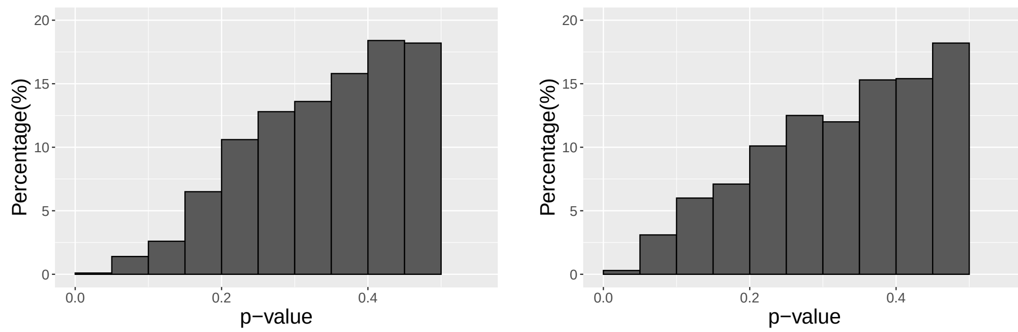
(b) Interference Test on Fairness

Figure 5 P-value Distribution of Interference Test (Human Group)

between 0 ad 0.5). This suggests that the productivity of workers in the human group is unlikely to have been affected by the interference between the algorithm group and the human group.

Figure 5(b) shows the distribution of p-values for the interference test on fairness across 1,000 draws of fixed subsets. The p-values from our 1,000 draws are smaller than 0.05 5.90% of the time, lower than the chance level of 10% under uniform distribution. This suggests that the perceived fairness of workers in the human group is unlikely to have been affected by the interference between the algorithm and human groups.

We next check the existence of interference for workers in the algorithm group. We follow the same steps as described above, except that we randomly select 12 workers from the algorithm condition as the fixed subset in each of the 1,000 draws. Figure 6(a) shows the distribution of p-values for the interference test on productivity across 1,000 draws of fixed subsets. Almost all p-values from the 1,000 draws are greater than 0.05 (99.90% of the time). Figure 6(b) shows the distribution of p-values for the interference test on fairness across 1,000 draws of fixed subsets. Again, almost all p-values from the 1,000 draws are greater than 0.05 (99.70% of the time). Thus,



(a) Interference Test on Productivity (b) Interference Test on Fairness

Figure 6 P-value Distribution of Interference Test (Algorithm Group)

we further confirm that the productivity and fairness perceptions of workers in the algorithm group are unlikely to have been affected by the interference between the algorithm group and the human group.

Online Appendix C: Second Field Experiment as Replication

Table 7 The Effects of Algorithmic (vs. Human-based) Assignment on Perceived Fairness and Productivity (Replication)

Dependent variable	*Standardized perceived fairness*			*Picking efficiency*		
	(1)	(2)	(3)	(4)	(5)	(6)
Algorithm	1.09**	1.13**	1.10**	1.21**	1.10**	1.01*
	(0.35)	(0.34)	(0.34)	(0.38)	(0.35)	(0.40)
Day fixed effects	No	Yes	Yes	No	Yes	Yes
Hour fixed effects	No	No	No	No	Yes	Yes
Demographics controls	No	No	Yes	No	No	Yes
Observations	87	87	87	3,181	3,181	3,181

Note: *p<0.05; **p<0.01; ***p<0.001. Average picking efficiency in the human group was 5.03.

We conducted another experiment from December 27th, 2019 to January 5th, 2020 to replicate the main results that we report in the paper. The second experiment involved 20 temporary picking workers. We randomly assigned them into either the algorithm group or the human group. Workers' experience with receiving and handling pick lists was the same as what we described in 4.1. These workers received 3,181 pick lists in total. The sample size of workers was smaller in the second experiment than the first experiment because (1) we could only run the second experiment for 10 days before a large sales period started on January 6th, 2020 and (2) the warehouse reduced labor floating, meaning that workers came to work for more days during the second experiment and consequently leaving us with fewer unique workers. Due to the small sample size of workers, we focus on replicating main effects, rather than heterogeneous treatment effects across different types of workers.

We distributed surveys at the end of each day to workers who worked at the warehouse that day. To assess their fairness perceptions, we asked workers two questions that contrasted algorithmic vs. human-based assignment processes. The first question asked workers, "Which assignment process do you think is more fair, algorithmic assignment or human-based assignment?" This question was measured on a five-point scale, with the anchors ranging from 1 ("Definitively algorithmic assignment") to 5 ("Definitively human-based assignment") for all workers. The second question was the same as that in the first experiment (see Table 1). For workers in the algorithm group, we reverse coded their answers to *both* questions; for workers in the human group, we made no adjustment to their original answers. Therefore, for workers in both groups, a higher (vs. lower) value on a question indicates that the worker viewed their current assignment process as more fair than the alternative process. The correlation between workers' responses to these two fairness questions (after reverse coding) was high ($r = 0.84$; $p < 0.0001$). Following the same procedure as described in 4.2, we created a score of *Standardized Perceived Fairness* for each worker each day. The survey also asked the same set of demographics as the survey described in the paper.

We analyze the effect of algorithmic (vs. human-based) assignment on fairness perceptions using specification (1) and report the results in Table 7 Columns 1-3. Column 1 (without any control variables) shows a positive and significant coefficient on the indicator *Algorithm* (p-value < 0.01), which indicates that receiving pick lists from an algorithm significantly increases workers' perceived fairness about their assignment process, compared to receiving pick lists from a human distributor. Specifically, algorithmic assignment (relative to human-based assignment) increases perceived fairness by 1.09 standard deviations. This effect is robust and even becomes slightly larger when we control for day fixed effects (1.13 standard deviations, p-value < 0.01; Column 2) as well as when we control for both day fixed effects and worker demographics (1.10 standard deviations, p-value < 0.01; Column 3). Overall, these results support Hypothesis 1 that assigning pick lists by an algorithm(vs. a human) boosts workers' perceived fairness about their assignment process.

We analyze the effect of algorithmic (vs. human-based) assignment on productivity using specification (2) and report the results in Table 7 Columns 4-6. Across all three columns with or without controls, the coefficient on the indicator *Algorithm* is positive and statistically significant (all p-values < 0.05), which means that the algorithmic assignment treatment significantly improves workers' productivity. Specifically, without control variables, we estimate that assigning pick lists via an algorithm increases worker productivity by 24.06%, relative to the average picking efficiency of 5.03 in the human-based assignment group (Column 4). When we control for day and hour fixed effects, the effect size decreases slightly: the percentage increase in productivity caused by

Table 8 IV-Estimated Effect of Perceived Fairness on Productivity (Replication)

Dependent variable	*Picking efficiency*		
	(1)	(2)	(3)
Standardized perceived fairness	1.20^{**}	1.04^{**}	0.99^{*}
	(0.38)	(0.33)	(0.40)
Day fixed effects	No	Yes	Yes
Hour fixed effects	No	Yes	Yes
Demographics controls	No	No	Yes
Observations	3,181	3,181	3,181

Note: *p<0.05; **p<0.01; ***p<0.001. Average picking efficiency across algorithm and human groups was 5.57.

algorithmic assignment (relative to human-based assignment) is 21.87% (Column 5). This effect is robust when we add demographics controls (20.08%; Column 6).

Table 8 shows the average treatment effect of perceived fairness on productivity using IV estimation based on specifications (3)-(4). We consistently find that workers' perceived fairness has a positive effect on productivity regardless of whether we include control variables (all p-values < 0.05 in Columns 1-3). Specifically, as perceived fairness increases by one standard deviation, worker productivity is estimated to significantly increase by 17.77%-21.54%, relative to the average picking efficiency of 5.57 across the algorithm and human groups (Columns 1-3).

Online Appendix D: Results of Interviews

In September, 2020, we conducted structured interviews with 13 picking workers (61.54% females, average age = 30.54) in the warehouse where our field experiments were implemented. Each interview lasted about 25 minutes on average. At the time of our interviews, hard-copy pick lists were printed and laid out on a table at the distribution station for workers to take. Note that the interviews took place one year after our first field experiment and eight months after our second field experiment. Considering that the workers in our field experiment were temporary workers, the workers in our interviews have a low chance of overlapping with workers in our field experiments. We could not verify this for sure since workers took our interviews anonymously and we could not match them with our field experiment data. In this online appendix, we summarized the key questions in order we asked workers, along with the key insights we gleaned from each question.

We first asked workers, "what factors usually influence your motivation and productivity?" The most frequently mentioned factors, brought up by 7 out of 13 workers, involve pick list characteristics including the number of items they have to collect and how many stocking positions they have to get products from. Another factor, which was brought up by 2 out of 13 workers, is the convenience of obtaining pick lists.

Next, we asked workers, "what factors could influence whether you find a pick list assignment process fair or unfair?" The most frequently mentioned factors, brought up by seven workers, involved the difficulty of assigned pick lists including the number of items they have to collect and how many stocking positions they have to get products from. Among these workers who mentioned task difficulty, most (N=5) also indicated that whether pick lists are assigned evenly across workers would affect their fairness perceptions as well. In addition, some workers (N=4) focused on the assignment process in the warehouse at the time of our interviews and complained that since pick lists were put on a table for workers to take, some of their colleagues tended to take pick lists according to their own preferences and leave harder pick lists to others, causing unfair task allocations.

Then we asked workers whether or not they thought the pick list assignment process would be fair if it was run by a human distributor as well as why they thought one way or the other. Among workers who indicated that a human-based assignment process might cause unfair outcomes (N=7), most (N=5) justified their evaluations by mentioning that they believed human distributors are subject to personal biases. For example, human distributors could give easier pick lists to workers who they personally know or who they have a good relationship with. Or workers could get difficult pick lists if they refuse to do personal favors for human distributors. We found out later that among workers who indicated that a human-based assignment process would be fair (N=6), two workers misunderstood our question. Specifically, they thought about human-based assignment as having workers take pick lists printed out by a human (i.e., the same as what was actually going on in their warehouse at the time of our interviews), rather than having a human allocate pick lists (i.e., what we were interested in knowing their thoughts about).

Furthermore, to get some sense about when workers care more fairness, we asked workers to rate how much they would care about the fairness of a pick list assignment process under three circumstances (from 1 = "Not at all" to 5 = "Very much"): their average response was 3.42 if they were paid based on the number of items they picked; their average response was 2.38 if they were paid by hour; and their average response was 4.00 if they were paid by their performance ranking among workers in the warehouse.

Next, we asked whether they thought the pick list assignment process would be more or less fair if they could receive pick lists by scanning a bar code than if they could receive pick lists from a human distributor. Most workers (N=10) believed the assignment process run by a machine would be more fair. When asked why they believed so, most workers (N=8) explained that they believed an algorithmic assignment process does not follow human distributors' personal preferences, would be able to deliver equal treatments across workers, and would not selectively favor or disadvantage certain workers.

In the end, we asked workers, "besides pick list characteristics and the assignment process, what other factors may influence your productivity?" Common factors brought up by workers include special circumstances (whether certain products are out of stocks, whether picking carts are temporarily unavailable), physical work environment (warehouse temperature, weather), and workers' physical well-being.

Online Appendix E: Supplement Results of Online Experiments

Using data from our online experiment reported in the paper, we test whether people who think equality should be prioritized over uniqueness (i.e., consideration of workers' personal characteristics) are more likely to find algorithms more fair. We split our sample by whether or not participants prioritize equality over uniqueness based on their separate importance ratings. Among participants who prioritize equality over uniqueness ($N = 123$), algorithmic assignment significantly increases (standardized) fairness perceptions ($M_{algorithm} = 4.37$, SD = 0.92), relative to human-based assignment ($M_{human} = 3.63$, SD = 1.03; t(114.96) = 4.19, p-value < 0.0001, Cohen's d = 0.77); among participants who prioritize uniqueness over equality($N = 44$) or view equality and uniqueness as equally important($N = 34$), the difference between conditions in fairness perceptions is directionally negative and not statistically significant ($M_{algorithm} = 3.97$, SD = 1.08 vs. $M_{human} = 4.00$, SD = 0.83; t(62.89) = 0.16, p-value = 0.88).

We also measured the relative importance of equality and uniqueness using one scale. We asked participants which of the two objectives they thought the warehouse should prioritize when it comes to assign picking tasks: treating all workers equality or taking into consideration personal characteristics. The anchors on the scale ranged from 1 ("Definitely should treat all workers equally") to 7 ("Definitely should consider workers' characteristics"). Choosing a higher (vs. lower) value indicates that the participant put less (more) weight on equality (uniqueness). Choosing the midpoint of the scale (i.e., 4) means that the participant thought it equally important to ensure equality and consider workers' unique characteristics. We first confirm that people on average prioritize equality over uniqueness in the warehouse task assignment setting ($M = 3.46 < 4$, $SD = 1.85$; t(200.00) = 26.55, p-value< 0.0001 for a one-sample t-test). Among participants who prioritize equality over uniqueness ($N = 108$), algorithmic assignment significantly increases (standardized) fairness perceptions ($M_{algorithm} = 4.45$, SD = 0.93), relative to human-based assignment ($M_{human} = 3.92$, SD = 1.06; t(100.06) = 2.78, p < 0.01, Cohen's d = 0.55); among participants who prioritize uniqueness over equality($N = 70$) or view equality and uniqueness as equally important($N = 23$), the difference between conditions in fairness perceptions is smaller and not statistically significant ($M_{algorithm} = 3.94$, SD = 1.01 vs. $M_{human} = 3.66$ SD = 0.85; t(82.35) = 1.42, p-value = 0.16).

We conducted another online experiment that followed the same design as the online experiment reported in the paper, except that we did not present information about pick list size in this additional experiment. A total of 200 participants from Amazon's Mechanical Turk comprised our study sample (38.50% female, $M_{age} = 38.925$). They were randomly assigned to either the algorithm condition (N=99) or the human condition (N=101). We replicate the results as follows:

Supporting the assumption underlying our Hypothesis 1, people view the assignment process run by a machine as more capable of preserving equality than the assignment process run by a human ($M_{algorithm} = 5.21$, $SD = 1.53$ vs. $M_{human} = 4.27$, $SD = 1.68$; t(196.91) = 4.17 p-value< 0.0001, Cohen's d=0.59). Further, in support of Hypothesis 1, participants in the algorithm condition perceived their assignment process more fair than those in the human condition ($M_{algorithm} = 3.56$, $SD = 0.87$ vs. $M_{human} = 2.98$, $SD = 1.04$; t(193.23) = 4.32, p-value < 0.0001, Cohen's d=0.61).

We test whether people who think equality should be prioritized over uniqueness are more likely to find algorithms fairer. We first split our sample by comparing participants' perceived importance of equality versus uniqueness on two separate measures. Among participants who prioritize equality over uniqueness ($N = 105$), algorithmic assignment significantly increases (standardized) fairness perceptions ($M_{algorithm}$ = 3.72, SD = 0.84), relative to human-based assignment (M_{human} = 2.82, SD = 1.09; t(95.69) = 4.66, p-value < 0.0001, Cohen's d = 0.92); among participants who prioritize uniqueness over equality($N = 56$) or view equality and uniqueness as equally important($N = 39$), the difference between conditions in fairness perceptions is much smaller and not statistically significant ($M_{algorithm}$ = 3.40, SD = 0.88 vs. M_{human} = 3.14, SD = 0.96; t(92.94) = 1.33, p-value = 0.19).

We next use the single item that assessed the relative importance of equality and uniqueness. Among participants who prioritize equality over uniqueness ($N = 92$), algorithmic assignment significantly increases (standardized) fairness perceptions ($M_{algorithm}$ = 3.83, SD = 0.88), relative to human-based assignment (M_{human} = 2.82, SD = 1.18; t(87.89) = 4.67, p-value < 0.0001, Cohen's d = 0.97); among participants who prioritize uniqueness over equality($N = 84$) or view equality and uniqueness as equally important($N = 24$), the difference between conditions in fairness perceptions is much smaller and not statistically significant ($M_{algorithm}$ = 3.36, SD = 0.81 vs. M_{human} = 3.13, SD = 0.87; t(103.81) = 1.43, p-value = 0.16).