­When humans and algorithms work together: Understanding perceptions of human-in-the-loop algorithmic decisions

| Min Kyung Lee  Machine Learning Department Carnegie Mellon University  Pittsburgh, USA  mklee@cs.cmu.edu | Stephanie Bao  Statistics & Data Science, and HCI Institute  Pittsburgh, USA  sbao@andrew.cmu.edu |
| --- | --- |

# ABSTRACT

Submission to CHI 2018

Algorithms increasingly automate or aid decisions that people used to make. How do people perceive these algorithmic decisions as compared to human decisions? We conducted a between-subjects online experiment with four managerial task scenarios to explore how perceptions of decision fairness, trust, and emotion vary by decision maker: human, algorithm, or human-in-the-loop (where humans and algorithms divide the functions of analyzing data and making decisions, in three configurations). For the assignment task requiring mechanical skills, decision maker had no effect. As tasks required more human skills, algorithmic and human-in-the-loop decisions were perceived increasingly more negatively than human decisions. For scheduling, the human analysis-based algorithmic decision was perceived as positively as the human decision in all aspects. For hiring, all human-in-the-loop decisions were perceived as more fair and trustworthy than algorithmic decisions, but less than human decisions. For evaluation, most human-in-the-loop decisions were perceived similarly to the algorithmic decisions. We discuss the social implications of algorithms.

## Author Keywords

Algorithmic management; human-in-the-loop; fairness; trust; emotion; decision-making.

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

# INTRODUCTION

Advances in artificial intelligence, machine learning and data infrastructure are transforming how people govern and manage citizens and organizations, changing the practices of managers, policy makers, physicians, teachers, police, judges, on-demand labor platforms, online communities, and more. Now more than ever, decisions that used to be made by people are being automated or aided by algorithms that process data and make data-driven decisions. Many of these decisions are embedded in governance or power structures, making them difficult to refute.

Algorithmic management directly impacts people’s lives, whether they are workers, students, patients, inmates, online community members, or even just residents of certain neighborhoods. How do people perceive this trend toward algorithmic management? Do people think decisions are more or less fair or trustworthy when algorithms vs humans make decisions? How do their perceptions change when humans and algorithms work together to make decisions? The answers will play a critical role in creating workplaces, communities, and societies that allow people to thrive.

Building on work that directly compared perceptions of algorithmic and human decisions, we conducted an online between-subjects experiment to compare perceptions of algorithmic and human decisions to perceptions of human-in-the-loop decisions, in which algorithms and humans divide the functions of analyzing data and making decisions. Three human-in-the-loop configurations were used: the human decision based on algorithmic analysis, the algorithmic decision using rules and factors monitored by humans, and the algorithmic decision based on human analysis. We focused on perceived fairness, trust, and emotional responses, key constructs that contribute to people’s satisfaction and social justice in organizations.

Our work makes the following contributions. The results offer a novel understanding of how people perceive hybrid decision-making involving both humans and algorithms, particularly when they are on the receiving end of such decisions, contributing to emerging theories on social aspects of algorithms. The results also offer insight into varying beliefs and concerns on algorithmic decisions, which can inform the communication and implementation aspects of human-in-the-loop systems in order to promote trust and a sense of fairness among those affected.

# Perception of human-in-the-loop algorithmic decisions

In this section, we first review the research on algorithmic and human decisions this paper builds upon. We then define human-in-the-loop algorithmic decisions and hypothesize how they may (or may not) differ from algorithmic and human decisions in perceived fairness, trust, and emotional responses.

## What are algorithms?

In this paper, we use the term “algorithm” to mean a process or set of rules to be followed in calculations or other problem-solving operations by a computer. Consistent with widespread use of machine learning and artificial intelligence and their portrayal in public media, we focus on algorithms’ capability to autonomously make decisions based on statistical models or decision rules without explicit human intervention. Treating algorithms as a sociotechnical system (cornell), emerging empirical studies explore how people think of and make sense of algorithms, especially in social media (Raedar, French, and others + cornellMooc). We investigate how algorithms and algorithmic decisions in management and governance contexts are viewed by the public.

## Perceptions of algorithmic vs human decisions depending on mechanical versus human tasks

Our research examines differences in perceptions of decisions depending on decision-maker: algorithm, human, or a combination of the two. Studies on source bias suggest that the same decisions can be perceived differently depending on perceivers’ knowledge of the decision-makers. Previous work on perceptions of other computational entities suggests two predictions. A long line of research referred to as “Computers Are Social Actors” has demonstrated that people may treat computers as social entities, applying human-human interaction principles to human-computer interaction, which suggests that algorithmic and human decisions might be similarly perceived. On the other hand, research on bots and robots suggest that people view them as having less agency and emotional capability than humans. Because the context of our inquiry does not involve direct interaction between algorithms and those affected by their decisions, we conjecture that different beliefs about the characteristics of algorithms and people, rather than interpersonal interaction principles, will inform people’s perceptions of algorithmic and human decisions.

Research done by Lee supports this hypothesis. Her experimental results suggest that human and algorithmic decisions are perceived differently depending on task characteristics—specifically, whether tasks require more mechanical or human skills. People thought algorithms could do more mechanical tasks as well as or better than humans, and perceived algorithmic and human decisions similarly in such contexts. However, people thought that humans could do more “human” tasks better than algorithms, and perceived human decisions as more fair, trustworthy, and positive in such contexts. We extend this framework in order to explore perceptions of human-in-the-loop algorithmic decisions.

## Human-in-the loop algorithmic decisions

The term “human-in-the-loop” refers to a configuration in which human control and/or supervision is part of an automated process. The term was first used in the field of human factors. Unlike closed-loop automation, human-in-the-loop systems allow people to intervene to correct errors and address ambiguity. Recently, the importance of having human-in-the-loop algorithmic decision systems for such functions as predictive policing, hiring platforms, and parole sentencing has been attracting attention. (xx) argues that for these algorithmic or predictive systems to be used in social, real-world settings, humans needs to be part of the algorithmic decision-making process. Below we describe three configurations of human-in-the-loop systems, where humans are involved in algorithmic input, the algorithm itself, or algorithmic output.

### Algorithmic analysis-based human decision

In the algorithmic analysis-based human decision configuration, algorithms analyze data and a human makes the decisions based on the analysis. Algorithms can process large amounts of data and narrow down or rank the options for the human decision-maker, reducing the human’s time and effort in decision-making. This configuration is the most commonly suggested interface in emerging algorithmic decision support applications.

### Algorithmic decision monitored by a human

In this configuration, a human monitors and adjusts the factors that algorithms use to make decisions. This configuration allows a human decision-maker to tailor algorithmic decision criteria for certain contexts, and intervene when algorithms produce less than ideal outcomes.

### Human analysis-based algorithmic decision

In the human analysis-based algorithmic decision, both humans and algorithms analyze the data and algorithms make decisions based on the analyses. This configuration has been used in decisions that require subjective analysis or more than one reviewer, such as essay grading. In this case, algorithms focus on quantitative data while humans focusing on qualitative data; this allows human decision-makers to focus on assessment, which algorithms cannot yet do reliably. The algorithm synthesizes the analyses to make the final decision.

## How would human-in-the loop situations influence perceived fairness, trust and emotional responses?

In this section, we hypothesize how perceptions of human-in-the-loop decisions might differ from those of algorithmic or human decisions depending in task characteristics.

### Fairness and Trust

Our paper focuses on perceived fairness and trust of decisions. Fairness means treating everyone impartially or equally (Leventhal, 1980). Perceived fairness in organizational decision has been shown to contribute to people’s sense of social justice and anti-organizational behaviors. In particular, we conjecture that different beliefs in decision-maker influence perceived procedural fairness, whether people think the procedures that regulate decision-making process fair. Because algorithms consistently follow the same procedure and are not influenced by interpersonal factors such as liking or favoritism, algorithmic decision might be perceived to have higher procedural fairness.

Trust is a belief and “attitude that an agent will help achieve an individual’s goals in a situation characterized by uncertainty and vulnerability”. Trust plays a central role in intraorganizational cooperation, coordination, and control (Kramer, 1999). In this paper, we focus on people’s trust in reliability and accuracy of decisions, and how much people believe in decision makers’ performance in the future. In automation technology, establishing the right level of trust has been shown to be challenging because of its unpredictable nature.

The research that compares algorithmic and human decisions suggests similar reasons/reasoning/factors contribute to perceptions of fairness and trust in algorithmic decisions. For mechanical tasks, algorithmic and human decisions did not differ in their perception, mainly because people believed both decisions makers are equally capable of making the decision, and there was little room for subjective judgment and potential human bias. Algorithmic decisions’ fairness and trustworthiness were attributed to its efficiency, objectivity in the process and those of human decisions were attributed to its managerial authority. Human-in-the-loop conditions involve both decision-makers, and both strengths can play complementary roles. Thus we hypothesize that all human-in-the-loop conditions will be perceived equally or even more fair and trustworthy than algorithmic or human decisions.

H1. For mechanical tasks, all human-in-the-loop conditions will be equally or even more fair and trustworthy than algorithmic or human decisions.

For human tasks, algorithmic decision was perceived less fair and trustworthy than human decision, mainly because people did not believe that algorithms were capable of making subjective judgments in tasks such as evaluating someone’s potential or performance. While algorithmic decisions were indeed perceived to have higher procedural fairness than human decisions, it did not mitigate the feeling of unfairness of using (incapable? Inadequate mechanics?).

People’s concerns with algorithms were the fact that algorithms would not be able to understand tones and emotions in conversation recording of a call center employees, or to judge someone’s potentials or interpersonal qualities from job applications. This assessment occurs when algorithms process data for decisions. Because the human analysis-based algorithmic decision has a human decision-maker analyze these data and uses as an input, we hypothesize that this configuration will be perceived fairer and more trustworthy than algorithmic decisions.

The algorithmic analysis-based human decision also has a human decision-maker making judgment, but after the algorithm’s analysis, and not everyone’s data will be reviewed by a human. The human-monitored algorithmic decision involves human decisions in the process, but a human decision maker’s role is to monitor and adjust factors that algorithms consider, not directly assessing data used by algorithms. This leads to our second hypothesis.

H2. For human tasks, all human-in-the-loop conditions will be perceived fairer and more trustworthy than algorithmic decisions, with the human analysis-based algorithmic decisions having greater effect than the other two (algorithmic analysis-based human decision and human monitored algorithmic decision).

### Emotional responses

Affective experience is important in job satisfaction and motivation. The algorithmic and human decision study suggests that for mechanical tasks, there was no difference in emotional responses. We hypothesize that human-in-the-loop decisions will also evoke similar emotional responses as algorithmic and human decisions.

H3. All human-in-the-loop decisions will evoke similar emotional responses as algorithmic and human decisions.

For human tasks, algorithmic decisions evoked more negative emotion than human decisions, because people felt dehumanized by being evaluated by machines and not by humans (dehumanization). All human-in-the-loop decisions involve human decisions in the process, and thus will evoke more positive emotional responses than algorithmic decisions. In particular, the human analysis-based algorithmic decision will have greater effect because the attention, time that human spend on individual cases will be great than the algorithmic analysis-based human decision and the human monitored algorithmic decision.

H4. All human-in-the-loop decisions will evoke more positive emotional responses than algorithmic decisions, with the human analysis-based algorithmic decision having greater effects than other two human-in-the-loop configurations.

**3. METHOD**

We conducted a between-subjects online experiment in July and August of 2017. Participants examined a managerial scenario, (Mintzberg, 1975) in which a decision was made by either humans, algorithms, or a combination of both. We examined how the type of decision-maker would influence people’s perceptions of the decisions by collecting both quantitative ratings of the decisions and qualitative reasons behind those ratings. We used a scenario-based method, which is commonly used in social psychology and ethics research, to investigate people’s opinions, beliefs, and attitudes (e.g., Petrinovich, O’Neill, and Jorgensen, 1993); studies have suggested consistency between people’s behaviors in scenario-based experiments and their behaviors in real life (Woods et al., 2006).

**3.1 Participants**

We recruited participants on Amazon Mechanical Turk (MTurk) to take an online survey that took 6.4 minutes on average to complete. Participants had to reside in the US, be at least 18 years old, have completed at least 100 HITs (Human Intelligence Task, MTurk’s task unit), and have at least a 95% HIT approval rate. Participants were compensated $1.00 for their time, a rate more than the minimum wage in the US ($7.25/hour). 1637 people responded. We omitted participants, indicated that they did not reside in the US (N=7), were younger than 18 (N=118), did not pass the first attention check (N=153), or did not finish the survey (N=7). These procedures left 1352 participants in the sample, about four-fifths of the original number. From these 1352 participants, we omitted those that failed the reading check (N=18) and failed the manipulation check (N=88). However, we allowed participants that failed the manipulation check but had correct explanations of the situation (N=18). We also omitted participants that correctly answered the manipulation check but failed to correctly explain the situation (N=90). In the end, there were 1174 participants in the sample.

Their ages ranged from 18 to 88, with an average age of 35.9 (SD = 11.39). 51% of the sample was female. They were fairly well educated: the mean education score was 4.12 (SD = 1.33; 3=”Some college or currently enrolled”, 4=”2-year degree”). The sample was 76% Caucasian, 8.28% Asian, 6.06% Hispanic, 7.42% African American, 0.43% Native American, 0.26% Pacific Islander, and 1.37% other.

**3.2 Materials**

We presented participants with four scenarios in which either algorithmic or human managers or a combination of both made decisions that had significant impacts on human workers (Table 1). Our scenarios were based on real-world situations in which algorithms are being employed in middle-management roles linked to the main functions of management (Mintzberg, 1975). We organized these four tasks based on the amount of “mechanical” skills and the amount of “human” skills involved, with assignment being the most mechanical and evaluation requiring the most human skills. The work assignment scenario included maintenance tasks in a factory, which used algorithms to predict the likeness of machinery components malfunctioning in order to prevent a larger stoppage in the overall workflow (Jardine, Lin, and Banjevic, 2006) (Table 1.a). The scheduling scenario used the example of an employee shift scheduling algorithm, which determined when café baristas would be called into work based on the predicted number of customers in the café at a given time and the baristas’ schedule preferences (Table 1.b) (Kantor, 2014; Pinedo, Zacharias, and Zhu, 2015).

|  |  |  |
| --- | --- | --- |
| **Managerial Decision Type** | | **General policy and scenario** |
| Tasks that require mechanical skills | Work Assignment | General policy: In a high-tech company, an/a [algorithm/ manager] analyzes personal statements and resumes on its job search website for each open engineering position. [A manager monitors and adjusts how the algorithm considers different factors in its decision.] The [algorithm/ manager] selects the top applicants for final on-site interviews.  Specific scenario: Chris applies for the engineering position on the job search website by submitting his personal statement and resume. The [algorithm/ manager] [, which is monitored and adjusted by the manager,] analyzes thousands of personal statements and resumes including Chris's. The [algorithm/ manager] selects the top applicants for final on-site interviews. |
|  | b. Work Scheduling | General policy: In a manufacturing factory, [algorithm/ manager] analyzes the importance of each component to the factory and how often they have worn off and broken down in the past. [A manager monitors and adjusts how the algorithm considers different factors in its decision.] The [algorithm/ manager] assigns the employees to check and update certain components of the machinery to prevent any critical operation failure.  Specific scenario: Chris works in the manufacturing factory. The [algorithm/ manager] [, which is monitored and adjusted by the manager,] analyzes the importance of the component to the factory and its breakdown history. The [algorithm/ manager] assigns him to check a specific component of the machinery and he does the maintenance work on it. |
| Tasks that require human skills | c. Hiring | General policy: In a customer service center, a/an [algorithm/ manager] analyzes the contents and tones of their calls with customers and their customer comments. [A manager monitors and adjusts how the algorithm considers different factors in its decision.] The [algorithm/ manager] evaluates the employees' performances.  Specific scenario: Chris works at the customer service center. The [algorithm/ manager] [, which is monitored and adjusted by the manager,] analyzes past call recordings and customer comments and the [algorithm/ manager] evaluates his performance. |
| d. Work Evaluation | General policy: In a coffee shop, a/an [algorithm/ manager] analyzes employee availability and the predicted number of customers at different times of the day to decide who should come into work and when. [A manager monitors and adjusts how the algorithm considers different factors in its decision.] The [algorithm/ manager] schedules the employees.  Specific scenario: Chris works in the coffee shop. The [algorithm/ manager] [, which is monitored and adjusted by the manager,] analyzes the employees' availabilities and the number of customers throughout the day. The [algorithm/ manager] schedules Chris to work mornings. |

**Table 1. Managerial scenarios presented to participants**

The hiring scenario was based on job search websites, such as LinkedIn, that use algorithms to analyze resumes and select top candidates for onsite interviews (Table 1.c). Finally, the work evaluation scenario revolved around a customer service call center that used a natural language-based algorithm to evaluate the performance of its employees (Petrushin, 1999) (Table 1.d).

We used the projective, third-person viewpoint while creating scenarios so that participants were reading scenarios that describe another person’s experience (e.g., Chris works […]) as supposed to ones that directly put the readers into the scenario (e.g., you work […]). The projective viewpoint has been shown to minimize social desirability effects, which is the desire to present socially desirable answers rather than honest opinions, and to have considerable external validity (Nisbett, et al., 1973). The scenarios had two parts. The first part described the general policy of management, and the second part described a specific instance of the policy involving a worker. We manipulated the type of decision-maker (algorithmic, human, or both) for all scenarios.

We conducted a pilot test to check which decision-makers people thought would perform each task better and whether those perceptions were in line with our assumptions; specifically, whether people thought the combination of the algorithm and the human would outperform the algorithm and/or the human alone. The within-subjects survey (N=21) presented the general policy portion of the four scenarios in a random order and asked which type of decision-maker would perform the task better (1=“Algorithms will perform the task better than human managers,” 2= “Both algorithms and human managers will perform the task equally well,” 3= “Human managers will perform the task better than algorithms”). We ran a multi-level analysis controlling for individuals for repeated measures. The results show that people thought that algorithms would perform the work assignment and scheduling tasks slightly better the human managers (Massign=1.4 (SE=.15), Mschedule=1.65 (SE=.15)), and that human managers would perform the hiring and evaluation tasks better than algorithms (Mhire=2.5 (SE=1.5), Mevaluate=2.9 (SE=1.5)).

**3.3 Procedure**

After receiving consent and affirming that the participants were over the age of 18 and a US resident, we gave the participants an attention check question adopted from Egelman and Peer, 2015. Failure to correctly answer the attention check resulted in immediate disqualification from the survey. Those who passed the check were randomly assigned to a decision-maker [human/algorithmic/both] in one of the four task scenarios. Participants in the algorithmic condition were shown this definition of “algorithm”: “Algorithms are processes or sets of rules that a computer follows in calculations or other problem-solving operations. In the situation below, an algorithm makes a decision autonomously without human intervention.”[[1]](#footnote-1) We presented the definition to ensure that participants understood our usage of the word algorithm, and bolded the analyst and the decision-maker of each scenario in order to emphasize any differences or similarities. Participants were then presented with a scenario and survey questions about their thoughts on the scenario they had just read. The manipulation check and demographic questions were asked at the end.

**3.4 Measures**

**3.4.1 Perceptions of decisions**

Except for a few open-response questions, all survey items used a 7-point Likert-type scale. Response options varied based on the questions (i.e., Strongly Disagree, Very Unfair).

**Decision fairness.** The question on the decision’s fairness was adopted from previous research (Brockner et al., 1994; Konovsky and Folger, 1991): “How fair or unfair is it for [scenario subject] that the [algorithm/manager takes the action specified in the scenario]?” For example: “How fair or unfair is it for Chris that the manager evaluates his performance? The scale ranged from “Very unfair” (1) to “Very fair” (7).

**Trust.** To ascertain subjects’ trust in the reliability and accuracy of the decision presented in each scenario, we asked, “How much do you trust that the [algorithm/manager] make good-quality [decision specified in the scenario]?” The scale ranged from “No trust at all” (1) to “Extreme trust” (7).

**Emotional response.** To understand how participants thought the decision affected the scenario’s subject, we asked how much they agreed or disagreed that the decision-maker’s decision would make the scenario’s subject feel happy, joyful, proud, disappointed, angry, and frustrated (1=“Strongly disagree”, 7=“Strongly agree”) (Larsson, 1987; Weiss, Suckow, and Cropanzano, 1999). We constructed an emotional response scale by averaging answers to the three positive adjectives and the reversed answers to the negative adjectives, so that greater numbers meant more positive emotion. The scale was very reliable (Cronbach’s α = .9).

**Open-response question.** After each of the three questions above, we asked participants to explain their reasons for their numeric ratings.

**3.4.2 Manipulation checks, attention checks, and demographic questions**

At the end of the survey in the algorithmic decision conditions, participants were asked an open-ended question: “In your own words, please briefly explain what you think algorithms are.” The answers confirmed that participants perceived algorithms as autonomous decision-makers. Another manipulation check question asked all participants: “Which of the following made the decisions in the situations that you read?” and provided a choice between humans, algorithms, or both. All participants correctly answered this question by condition.

We used one attention check (Egelman and Peer, 2015) in the beginning of the survey to immediately disqualify participants. Throughout the survey, we also asked participants how much they agreed with the statement, “I do not read the questions in this survey.” At the end of the survey, we also asked participants to indicate their knowledge of algorithms (1=”No knowledge at all”, 5=“Expert knowledge of algorithms”) along with the demographic questions.

**3.5 Analysis**

We conducted a one-way ANOVA to find the main effect of the decision-maker on perception of the decision for each decision scenario and a multi-level analysis on the main and interaction effects of the decision-maker type and task types. We qualitatively analyzed participants’ reasons for their answers to the three questions about fairness, trust, and emotional response (Strauss and Corbin, 1990). We open-coded data at the response level in conjunction with participants’ survey ratings to identify emerging themes. We grouped different themes to explain how participants responded to and judged the decisions depending on the type of decision-maker.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Managerial Decision Type** | | **Decision-maker Type** | | | | | |
| **Measures** | **Algorithm** | **Human decision based on algorithmic analysis** | **Algorithmic decision (whose rule is monitored by a human)** | **Algorithmic decision based on both human and algorithmic analyses** | **Human** |
| Tasks ordered from involving least to most human skill | a. Work Assignment | Fairness | 5.57 (SE=.15) | 5.92 (SE=.15) | 5.76 (SE=.17) | 5.79 (SE=.16) | 6.03 (SE=.15) |
| Trust | 5.21 (SE=.14) | 5.35 (SE=.14) | 5.50 (SE=.15) | 5.52 (SE=.14) | 5.13 (SE=.14) |
| Emotion | 4.70 (SE=.12) | 4.76 (SE=.13) | 4.41 (SE=.14) | 4.65 (SE=.13) | 4.63(SE=.12) |
| b. Work Scheduling | Fairness | 4.97 (SE=.16) | 5.41 (SE=.16) | 5.19 (SE=.17) \*H | 5.77(SE=.19) \*A | 5.71 (SE=.15) |
| Trust | 4.31 (SE=.15) | 4.76 (SE=.15) | 4.66 (SE=.17) \*H | 5.18 (SE=.19) \*A | 5.18 (SE=.15) |
| Emotion | 4.15 (SE=.13) | 4.47 (SE=.13) \*H | 4.21 (SE=.14) \*H | 4.77 (SE=.16) | 4.86 (SE=.13) |
| c. Hiring | Fairness | 4.37 (SE=.17) | 4.89 (SE=.19) \*A, \*H | 4.98 (SE=.19) \*A, \*H | 5.04 (SE=.20) \*A, \*H | 6.17 (SE=.17) |
| Trust | 3.93 (SE=.17) | 4.67 (SE=.18) \*A, \*H | 4.60 (SE=.19) \*A, \*H | 4.48 (SE=.20) \*A, \*H | 5.48 (SE=.17) |
| Emotion | 3.69 (SE=.16) | 3.94 (SE=.17) \*H | 3.93 (SE=.18) \*H | 4.01 (SE=.19) \*H | 4.62 (SE=.16) |
| d. Work Evaluation | Fairness | 3.86 (SE=.18) | 3.77 (SE=.19) \*H | 4.28 (SE=.19) \*H | 3.96 (SE=.21) \*H | 6.08 (SE=.19) |
| Trust | 3.16 (SE=.19) | 3.38 (SE=.20) \*H | 3.63 (SE=.19) \*A, \*H | 3.62 (SE=.22) \*A, \*H | 4.94 (SE=.19) |
| Emotion | 3.28 (SE=.14) | 3.32 (SE=.15) \*H | 3.49 (SE=.15) \*H | 3.69 (SE=.17) \*A, \*H | 4.51 (SE=.15) |

**Table 2. People’s perceptions of algorithmic vs HitL vs human managerial decisions**

|  |  |  |  |
| --- | --- | --- | --- |
| General Definition | Algorithm/Human/HitL | Fairness/Trust/Emotion | Definition |
| No human bias in algorithm | Algorithm only | F, T, E | Algorithms are not biased like humans, in ways such as personal feelings and favoritism. |
| Efficiency in algorithm | Algorithm only | F, T, E | Algorithms efficiently make decisions, saving time and eliminating human mistakes/errors. |
| General trust in algorithm | Algorithm only | F, T, E | There is general trust in the algorithm's ability to complete the tasks accurately. |
| Lack of human judgment in algorithms | Algorithm only | F, T, E | Algorithms cannot sufficiently analyze some factors in place of human.  For example, they cannot account for edge cases, subjective or unmeasurable data, social chemistry/work culture.  General disapproval of using algorithms because the task requires human touch/judgement. |
| General concern about algorithmic error (distrust in algorithms) | Algorithm only | F, T, E | There is general distrust in the algorithms' ability to complete the tasks.  For example, algorithms cannot analyze certain data such as tones, personality, nuances, possible human mistakes, etc. There may be potential bugs in the algorithms. Also, just any general disapproval of using computers. |
| Positive combination of algorithm and human | HitL | F, T, E | Algorithms and human managers work together to successfully complete the tasks |
| Negative combination of algorithm and human | HitL | F, T, E | Having both algorithm and human in the decision-making process compromise the effectiveness of one decision maker or the other  For example, human involvement can compromise algorithm's objectivity; human input will not be fairly considered by the algorithm |
| Human bias | HitL, Human | F, T, E | Humans are inherently biased. |
| Equal treatment in the decision-making process | Both | F, T, E | Everyone is considered in the same manner/process. |
| Managerial authority | Both | F, T, E | The task is under the company's and manager's authority.  There is general trust in who programmed the algorithm. |
| Use of (good/trustworthy) data source for decision-making | Both | F, T, E | Decisions are made based on trustworthy data. |
| Employee/job contract | Both | F, T, E | The task is part of the employee’s job.  There may be general work practice issues. |
| Fairness concept is not applicable | Both | F, T, E | The fairness/trust/emotional concept does not apply.  For example, because the decision is independent of the decision maker, perhaps because it is based on the data)  The algorithm is based on luck/chance |
| Needs more info | Both | F, T, E | The mTurker expressed that they need more information on what happened to make a decision. |
| Nonsense, irrelevant, taken personally by mTurker | Both | F, T, E | Not enough information is provided in the participants' answers (not explicit in explaining their answers).  Irrelevant (researchers cannot understand the connection). |
| Chris feels \_\_\_\_ because of \_\_\_ factor(s) | Both | T, E | The mTurker’s answer is dependent on the result (i.e. Chris was hired).  The mTurker’s answer is dependent on his or her own personal preferences. |
| neutral feeling about being judged by algorithms/machines | Algorithm, HitL | T, E | Neutral. |
| General positive emotion | Both | T, E | General positive emotion toward the situation. |
| General negative emotion | Both | T, E | General negative emotion toward the situation. |

1. Oxford Dictionaries http://www.oxforddictionaries.com/ [↑](#footnote-ref-1)