

CONFIDENTIAL - FOR PEER-REVIEW ONLY

The Effect of Algorithmic Management on Work Performance (#122067)

Created: 02/15/2023 04:27 AM (PT)

This is an anonymized copy (without author names) of the pre-registration. It was created by the author(s) to use during peer-review. A non-anonymized version (containing author names) should be made available by the authors when the work it supports is made public.

1) Have any data been collected for this study already?

No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?

The main question being asked in this study is whether there is a difference in employees' work performance under algorithmic management compared to human management. We argue that the answer to this question largely depends on the manager's role and relationship with the employee, relative to other working conditions, including factors such as salary, job description, and required skills. If material factors predominantly determine worker behavior, we hypothesize that algorithmic management will have only a negligible effect. On the other hand, if the relationship between workers and managers is a key motivating factor, as other studies suggest, we expect to see a significant difference in the work of those supervised by an algorithmic manager compared to those by a human manager. This difference is likely to be influenced by the nature of the interaction that workers have with their managers.

1. Effort and performance.

* H1a: People will put less effort into work when experiencing a negative interaction with an algorithmic manager than with a human manager.

* H1b: People will put more effort into work when experiencing a positive interaction with a human manager than with an algorithmic manager.

2. Commitment:

* H2: The impersonal feature attributed to algorithms will make the relationship between the manager less binding, reducing emotional commitment to the employer, especially when decisions made by the manager are unfavorable.

3. Work satisfaction and acceptance:

H3a: In the absence of an emotional connection with their manager, employees may complete their work without attaching any significant positive or negative meaning to it due to the impersonal nature of algorithmic management.

H3b: In the presence of the emotional component, a negative interaction with a human manager will result in a greater decline of employee satisfaction compared to a negative interaction with an algorithmic manager.

3) Describe the key dependent variable(s) specifying how they will be measured.

We will measure the following behavioral aspects of work:

Our primary measure of work performance will be the number of errors made during a task that involves evaluating eight open-ended comments on the future of fashion. Participants will rate each comment on a scale from very negative (-100) to very positive (+100) without considering their personal views. The task will consist of five comments with predictions having a positive or negative tone (randomly selected), two neutral comments, and one comment with the opposite tone to the other comments. We will consider a response as an error if the respondent fails to place a neutral comment within 10% of each direction from the 0-point on the scale, or if the respondent fails to classify the comment with the opposite tone as closer to the end of the scale than the other comments. For secondary outcomes, we will use (1) a measure of the time it takes workers to complete the task, and (2) a measure of the number of clicks as an indicator of revisions.

To measure commitment, we will ask: "Would you like our algorithm/team member to consider you for additional tasks? If so, at what wage?" For each task (rating and cataloging), we will calculate the median requested wage and measure the difference between the median and the requested wage, with a negative value (-) indicating high commitment and a positive value (+) indicating low commitment.

To measure work satisfaction, we will ask workers after completing the task, 'How much did you enjoy the task assigned by the algorithm?' with answers on a scale of 1 to 10, where higher values represent more satisfaction. For the secondary outcome, we will use answers to the question 'How fair do you think the pay was for the task you performed?' on a scale of 1 to 5.

To assess the acceptance of management authority, we will ask the following question: "After completing the cataloging/rating task, how well do you understand why the algorithm/team member assigned you to it?" using a 1-4 point scale, where a higher value indicates a better understanding.

4) How many and which conditions will participants be assigned to?

To evaluate the impact of the manager's identity on worker behavior, the main intervention alters the decision-maker who hires and assigns workers to tasks. The treatment group will be informed that a computer algorithm was used to assign workers to tasks, while the control group will be told that an HR team member made the assignments. The second intervention distinguishes between a positive and negative experience with the manager by varying the description of the task, either a high-paying, higher-status task or a low-paying, lower-status task. We will exclude participants who prefer the lower-status task, resulting in two treatment groups: those with a positive experience assigned to the high-paying-status desired task and those with a negative experience assigned to the low-paying-status undesired task.

As part of another, unrelated study that assesses the effect of exposure to information on shaping attitudes, we will randomly vary the tone of the comments used to assign tasks to workers (positive versus negative tone). We plan to combine these two groups, as we do not expect significant

differences in the outcome of interest between these two subgroups.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

Our main target parameter is the intent-to-treat effect. We will use ordinary least squares to estimate the average treatment effect (ATE) and adjust for pre-treatment covariates, such as Age, Gender, Education, Race, Party ID, and Digital Literacy, to obtain a more precise estimate of the treatment effect. We will use the linear regression model below to test the hypotheses, varying the outcome of interest: $Y_i = \alpha_i + \beta_{\text{employer}_i} + \beta_{\text{task}_i} + \beta_{\text{employer}_i \times \text{task}_i} + \beta X_i + \epsilon_i$ Here, β_{employer} is the coefficient of the employer treatment (with the human employer as the reference category), β_{task} is the coefficient of the type of experience with the employer (with the negative interaction, which is defined by assignment to the undesired task as the reference category), $\beta_{\text{employer} \times \text{task}}$ is the interaction term, X is a matrix of demographic covariates, and ϵ_i is the error term. The dependent variable Y_i can be one of three possible variables, depending on the hypotheses: the number of errors (0-3), the gap between the wage the participant proposed upon completing a similar assignment under the same manager minus the median proposed wage for the task assigned to, or the participant's answer to the satisfaction question (on a scale of 1-10).

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

The target population is MTurk workers, that is, individuals paid to complete a series of Human Intelligence Tasks (HITs). Participants who do not answer correctly on the attention test at the beginning of the pretreatment survey will be excluded from the study. We will exclude the top 1% of the most active workers on MTurk, who complete around 21% of the daily HITs on the platform, to minimize the potential impact of familiarity with AI on our results and to enhance the external validity of our findings to other labor markets. As part of another unrelated study, in which we empirically assess the effect of exposure to new information, negative or positive, concerning artificial intelligence on AI-based attitudes, we utilize the same experiment setting, manipulating the content of the task that participants perform. We will limit the sample to include only subjects exposed to the placebo information about the future of fashion (negative and positive).

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.

Assuming a power of 0.8 and an alpha level of 0.05 (two-tailed), to detect small effects of 0.2 and assess interaction hypotheses, power calculations suggest a sample of 40 participants in each of the 8 cells, resulting in a total sample size of at least 640 participants. To account for a 40% attrition rate for each wave (wave 2 and wave 3) and ensure a final sample size of 640 participants, we need to recruit at least 1300 participants.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

To ensure treatment validity, we will ask participants to rate their satisfaction with the task assigned by the algorithmic or human manager on a 5-point scale. We will also include an open-ended question for participants to explain why they think the decision was wrong.

BUNDLE

This pre-registration is part of a set of similar and/or related pre-registrations sharing at least one author. When one of these pre-registrations was shared by an author, the rest were shared automatically. Links to other pre-registration(s), appear below:

#121928 - https://aspredicted.org/XNX_KCB - Title: 'The Evolving Politics of using AI-based Algorithms in Public Policy'