

Trust in Management Automation over Time (#33244)

Created: 12/20/2019 06:32 PM (PT)

Public: 06/24/2022 02:02 AM (PT)

Author(s)

Markus Langer (Universität des Saarlandes) - markus.langer@uni-saarland.de

Victoria Helsing (Universität des Saarlandes) - helsingv@gmail.com

Caroline Back (Universität des Saarlandes) - carolineback@aol.com

Cornelius J. König (Universität des Saarlandes) - ckoenig@mx.uni-saarland.de

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?

a) There is still little research on how trust in algorithms or support by an automated-system for management decisions changes over time and how automated decision-support systems affect users. This study compares 4 conditions. In a personnel selection scenario, participants will get a preselection of applicants either from an automated system or from a human colleague. Additionally, participants either receive information that the system/colleague is imperfect or they receive no information about the performance of the decision-support agent. Each participant completes twelve situations. We will test how trust processes over time (trust building, trust violation, trust repair) differ for a preselection provided by an automated system or a human colleague. The leading research question for this study therefore is: Do trust processes towards human decision-support differ from trust processes towards automated decision-support?

Hypotheses:

1. When registering an error, Trustworthiness, Trust and Reliance towards automated support decreases to a larger extent compared to human support.
2. After the error has been corrected and after there has been an excuse, Trustworthiness, Trust and Reliance in automated support increases to a lesser extent compared to human support.
3. Trustworthiness, Trust and Reliance in imperfect automated support will initially be lower compared to automated support without information.
4. Trustworthiness, Trust and Reliance in imperfect automated support will decrease less when registering a failure compared to automated support without information

Exploratory

1. The acceptance/rejection of human support is assessed as deontological.
2. The acceptance/rejection of automated support is evaluated as consequential.
3. The acceptance/rejection of imperfect automated support is evaluated as more deontological than the acceptance/rejection of automation without information.

Research Questions:

1. Is there an initial difference between trustworthiness, trust, and reliance in regard to human vs. automated support?
2. Will participants realize trust violations by automated decision support later? (e.g., because they expect the system to behave in an utilitarian way and to not be biased)
3. Is there an interaction between the support-agent and knowledge of imperfection of the support agent regarding the trust processes? (i.e., will information regarding imperfection only affect the automated system?)
4. Is there a difference on the effects of trust violations and trust repair on the facets of trustworthiness (e.g., will the facet ability suffer more than the other facets?), the evaluation of trust or reliance?
5. Will knowledge about imperfection affect moral reasoning about rejection/acceptance of automation?

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

- a. Trustworthiness, including the facets ability, flexibility, benevolence, and integrity
- b. Intention to trust
- c. Propensity to trust
- d. Reliance, measured as the rejection/acceptance of decision-support recommendations as well as active supervision of decision-support. Rejecting a recommendation would indicate less reliance; supervision of the decision-support system would indicate less reliance
- e. Moral reasoning about the acceptance/rejection of support

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

a) The experiment will be conducted in the laboratory, except for when participants have no opportunity to get to the laboratory. In a 2 (algorithm vs. human support) x 2 (no information regarding the support vs. information regarding imperfection) participants will be randomly assigned to one of the four conditions. Participants should imagine being a hiring manager tasked with 12 consecutive personnel selection tasks. They are informed that they receive support by either a human or an automated decision-support agent. In the case of the imperfect decision-support groups, participants receive information that the decision-support agent is imperfect.

For every task round, the decision-support agent provides participants with a preselection of applicants. Specifically, participants will receive information about male and female applicants. After each preselection, participants state if the preselection followed the goals of the company (e.g. diversity). Participants then respond to questions regarding trustworthiness (tasks 1,3,5,7,9,12) of the algorithm or the human support and the intention to trust the respective support agent (tasks 1-12). Furthermore, participants will be asked if they want to see additional statistical information regarding the applicant pool for a given selection round (i.e., this should capture if participants rely on the decision-support without supervising it). In the end of every round, participants indicate if they would accept or reject the recommendation. In three rounds, participants will additionally respond to items assessing their moral reasoning when accepting or rejecting a recommendation (tasks 3,7,11).

Beginning in round 5, the preselection will predominantly select male applicants in order to decrease diversity in the preselection (i.e., this should be a trust violation, as the support system violates the goals of the company). This will continue until task 8. In the beginning of task 9, participants are informed that there was an issue with the preselection that is now solved (i.e., an excuse for the imbalanced preselection in tasks 5 - 8). The rest of the preselection rounds are balanced again. In the end, participants respond to questions about their perceptions regarding the support (algorithm/colleague) and whether they would recommend the support to others. Furthermore, participants answer questions on demographic information. Finally, they answer questions related to their prior experience with automation and application procedures and to questions assessing propensity to trust in humans or automation.

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

For the analyses of hypotheses and research questions, we will use either mixed ANOVA or hierarchical linear models. This will depend on the fulfilment of preconditions for the given statistical analyses.

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

We will exclude participants who paused the experiment for more than one hour, participants who showed insufficient effort in responding (i.e., quickly skimming through the provided information), participants indicating that their data should not be used, and participants with serious technical issues.

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.

For a within-between effect size of partial eta squared .01, a power of .80, an alpha-error of .05 we would need about 108 participants. Assuming a between effect of .04 we would need 148 participants for a power of .80.

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

We might also consider using time series analyses for testing of the hypotheses and research questions.