Study Information

Title (required)

Understanding Medical Algorithm Phenomenon – Part 2

Authors (required)

Matthew Leitao

Description (optional)

This study aims to assess differences between trust in Artificial Intelligence and Medical Doctors across various types of medical decisions. A secondary objective is to determine if quantitative reasoning and intuition are related to the constructs of warmth and competence.

Hypotheses (required)

H1. People will report trusting doctors more than artificial intelligence for all the medical decisions that we present to them.
H2. The difference in trust between doctors and artificial intelligence will vary based on the perceived amount of needed quantitative reasoning believed to be needed to make that medical decision.
H3.a. The capability of a decision-maker to use quantitative reasoning will correlate with the perceived competence of the decision-maker.
H3.b. The capability of a decision-maker to use intuition will correlate with the perceived warmth of the decision-maker.
H4.a. The perceived warmth of the decision-maker will mediate the differences in trust between the artificial intelligence and the doctor.
H4.b. The perceived competence of the decision maker will mediate the differences in trust between the artificial intelligence and the doctor.

Design Plan

In this section, you will be asked to describe the overall design of your study. Remember that this research plan is designed to register a single study, so if you have multiple experimental designs, please complete a separate preregistration.

Study type (required)

Choose one below

1) Observational Study - Data is collected from study subjects that are not randomly assigned to a treatment. This includes surveys, natural experiments, I and regression discontinuity designs.

Blinding (required)

Blinding describes who is aware of the experimental manipulations within a study. Mark all that apply.

Choose one below

1) No blinding is involved in this study.

Is there any additional blinding in this study?

N/A

Study design (required)

This study is a within-subjects, repeated measures survey where we are asking participants to make judgments about various medical decisions and various medical decision-makers. We will utilize both paired samples t-tests and linear mixed effects modeling to determine the differences between the judgments.

Randomization (optional)

There is no participant randomization occurring within this study.

Sampling Plan

In this section we'll ask you to describe how you plan to collect samples, as well as the number of samples you plan to collect and your rationale for this decision. Please keep in mind that the data described in this section should be the actual data used for analysis, so if you are using a subset of a larger dataset, please describe the subset that will actually be used in your study.

Existing data (required)

Preregistration is designed to make clear the distinction between confirmatory tests, specified prior to seeing the data, and exploratory analyses conducted after observing the data. Therefore, creating a research plan in which existing data will be used presents unique challenges. Please select the description that best describes your situation. Please do not hesitate to contact us if you have questions about how to answer this question (prereg@cos.io).

Choose one below

1) Registration prior to creation of data: As of the date of submission of this research plan for preregistration, the data have not yet been collected, created, or realized.

Explanation of existing data (optional)

N/A

Data collection procedures (required)

We intend to recruit participants using the Cloud Research Platform's Prime Panel feature. The study will take approximately 20 minutes and participants will be compensated $1.75 for their time.
Once recruited, participants will be sent to complete a Qualtrics survey where they will be asked about their judgments concerning medical decision-making and medical decision-makers before filling out some personal demographics.

Sample size (required)

103

Sample size rationale (optional)

For this study, we used half the smallest Cohen's d effect between the general doctors and the general algorithm that we found in our previous study. This smallest difference was for the difference in quantitative reasoning, which had a Cohen's d = .56.
.56/2 = .28.

```
t tests-Means: Difference between two dependent means (matched pairs)

Analysis:     A priori: Compute required sample size
Input:        Tail(s)                          =      Two
```

	Effect size dz	=	.28
	α err prob	=	0.05
	Power (1−β err prob)	=	0.8
Output:	Noncentrality parameter δ	=	2.8416896
	Critical t	=	1.9834953
	Df	=	102
	Total sample size	=	103
	Actual power	=	0.8036905

Stopping rule (optional)

N/A

Variables

In this section you can describe all variables (both manipulated and measured variables) that will later be used in your confirmatory analysis plan. In your analysis plan, you will have the opportunity to describe how each variable will be used. If you have variables which you are measuring for exploratory analyses, you are not required to list them, though you are permitted to do so.

Manipulated variables (optional)

N/A

Measured variables (required)

Our outcomes are linked to the following scales.

The first scale measures the level of trust people would place in either a Doctor or an Artificial Intelligence across 12 different medical decisions. This scale goes from 0 (Don't Trust at All) to 10 (Completely Trust).

The twelve decisions:
Recommend stretching to prevent injury
Diagnosing a broken bone
Prescribing pain medication for a knee injury
Recommend meditation and exercise to promote mental health
Diagnose depression
Prescribe anti-depressants and cognative behavioral therapy for depression
Recommend vitamin c + zinc to fight off infections
Diagnose strep throat
Prescribe anti-biotics for strep throat
Recommend a strict diet to promote better physical health
Diagnose diabetes
Prescribe insulin and a change in diet for diabetes

The second scale measures participants' perceived percentage of quantitative reasoning (versus intuition) required by medical professionals to make accurate medical decisions for the previous thirteen medical decisions. This scale goes from 0 (Requires only Intuition and Personal Experience) to 100 (Requires only Quantitative Reasoning).

The third set of questions pertains to the perceptions of Doctors and Artificial Intelligence.

Here are the set of questions:

1) From 0-10, with 0 being "Not Capable at All" to 10 being "Completely Capable", how capable is each entity of using quantitative reasoning to make recommendations about health?
2) From 0-10, with 0 being "Not Capable at All" to 10 being "Completely Capable", how capable is each entity of using intuition to make recommendations about health?

3) From 0-10, with 0 being "Not Capable at All" to 10 being "Completely Capable", how capable is each entity at making accurate recommendations about health in general?
4) From 0-10, with 0 being "Do Not Believe at All" to 10 being "Completely Believe", how much do you believe that each entity can anticipate your future health needs?
5) From 0-10, with 0 being "Do not Believe at All" to 10 being "Completely Believe", how much do you believe that each entity understand your feelings surrounding your health?
6) From 0-10, with 0 being "Do not Believe at All" to 10 being "Completely Believe", how much do you believe that each entity can account for your health needs?
7) From 0-10, with 0 being "Not Familiar At All" to 10 being "Completely Familiar", how much does each entity's decision making process reminds you of receiving other types of health recommendations?
8) From 0-10, with 0 being "No Transparent at All" to 10 being "Completely Transparent", how transparent is decision making process for each entity?
9) From 0-10, with 0 being "Do Not Understand at All" to 10 being "Completely Understand", how much do you understand the process of which each entity comes up with a health recommendation?
10) How much do you believe competent and capable describes the entity?
11) How much do you believe warm and friendly describes the entity?

Indices (optional)

To determine the differences between preventative, diagnostic, and prescriptive medical decisions, we will create an average of the within-person responses for each category of medical decision.

Diagnostic Decisions:

1. Diagnosing a Broken Bone
2. Diagnosing Depression
3. Diagnosing Strep Throat
4. Diagnosing Diabetes

Prescriptive Decisions:

1. Prescribing Pain Medication for a Knee Injury
2. Prescribing Anti-Depressants and Cognitive Behavioral Therapy for Depression
3. Prescribing Anti-Biotics for Strep Throat
4. Prescribing Insulin and a Change in Diet for Diabetes

Preventative Decisions:

1. Recommending Stretching to Prevent Injury
2. Recommending Meditation and Exercise to Promote Mental Health
3. Recommending Vitamin C + Zinc to Fight Off Infections
4. Recommending a Strict Diet to Promote Better Physical Healt

To determine the differences between decisions made about bodily injuries, mental health, acute sickness, and chronic sickness, we will create an average of the within-person responses for each category of medical decision.

Bodily Injury:

1. Recommending Stretching to Prevent Injury

2. Diagnosing a Broken Bone
3. Prescribing Pain Medication for a Knee Injury

Mental Health:
1. Recommending Meditation and Exercise to Promote Mental Health
2. Diagnosing Depression
3. Prescribing Anti-Depressants and Cognitive Behavioral Therapy for Depression

Acute Sickness:
1. Recommending Vitamin C + Zinc to Fight Off Infections
2. Diagnosing Strep Throat
3. Prescribing Anti-Biotics for Strep Throat

Chronic Sickness:
1. Recommending a Strict Diet to Promote Better Physical Health
2. Diagnosing Diabetes
3. Prescribing Insulin and a Change in Diet for Diabetes

If participants are missing any of these values, we will take the average of the remaining responses for the category.

We will create a mind perception composite by averaging across these seven questions. If participants are missing any of these values, we will take the average of the remaining responses for the category.

1) From 0-10, with 0 being "Not Capable at All" to 10 being "Completely Capable", how capable is each entity at making accurate recommendations about health in general?
2) From 0-10, with 0 being "Do Not Believe at All" to 10 being "Completely Believe", how much do you believe that each entity can anticipate your future health needs?
3) From 0-10, with 0 being "Do not Believe at All" to 10 being "Completely Believe", how much do you believe that each entity understand your feelings surrounding your health?
4) From 0-10, with 0 being "Do not Believe at All" to 10 being "Completely Believe", how much do you believe that each entity can account for your health needs?
5) From 0-10, with 0 being "Not Familiar At All" to 10 being "Completely Familiar", how much does each entity's decision making process reminds you of receiving other types of health recommendations?
6) From 0-10, with 0 being "No Transparent at All" to 10 being "Completely Transparent", how transparent is decision making process for each entity?
7) From 0-10, with 0 being "Do Not Understand at All" to 10 being "Completely Understand", how much do you understand the process of which each entity comes up with a health recommendation?

Analysis Plan

You may describe one or more confirmatory analysis in this preregistration. Please remember that all analyses specified below must be reported in the final article, and any additional analyses must be noted as exploratory or hypothesis-generating.

A confirmatory analysis plan must state up front which variables are predictors (independent) and which are the outcomes (dependent), otherwise, it is an exploratory analysis. You are allowed to describe any exploratory work here, but a clear confirmatory analysis is required.

Statistical models (required)

Our primary analysis will be a paired samples t-test to compare trust in doctors and artificial intelligence along every dimension of medical decisions. We will also compare doctors and artificial intelligence using paired samples t-tests on the medical decision dimensions we will create.

We will also conduct a linear mixed-effect model to investigate the overall trust difference in doctors and artificial intelligence for all medical decisions without the composites. Next, we will run this linear mixed-effects model using people's perceptions of quantitative reasoning/intuition needed to make accurate medical decisions.

Additionally, we will run a linear mixed-effect model to determine whether the perceived warmth and competence of the decision-maker mediate the difference in trust between doctors and artificial intelligence.

Finally, we will correlate mind perception, the capability to use intuition, the capability to use quantitative reasoning, competence, and warmth.

Transformations (optional)

N/A

Inference criteria (optional)

We will be using the alpha = .05 criterion to determine significance.

Data exclusion (optional)

We don't intend to exclude any participants.

Missing data (optional)

We intend to include all data from which participants have answered all the questions for our dependent variables.

Exploratory analysis (optional)

We will use add other covariates to our model, including age, gender, education, and religiosity.

We will do a person-effect-size analysis for the differences in trust between doctors and artificial intelligence.

Other

Other (Optional)

None.