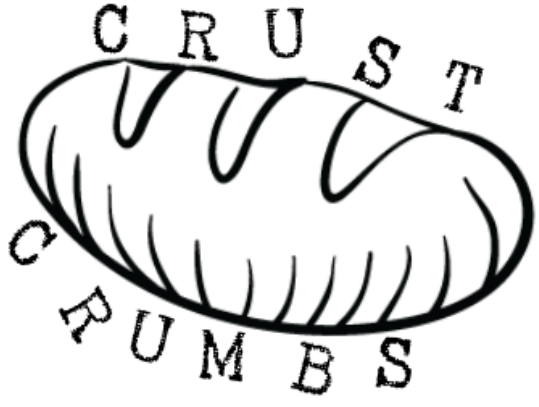
Reproducibility

Theory, Social Aspects, and Practice Of

Team: L. Gustavo Nardin, Berna Devezer, Bert Baumgaertner, Erkan Buzbas

Updated: 2017/09/26

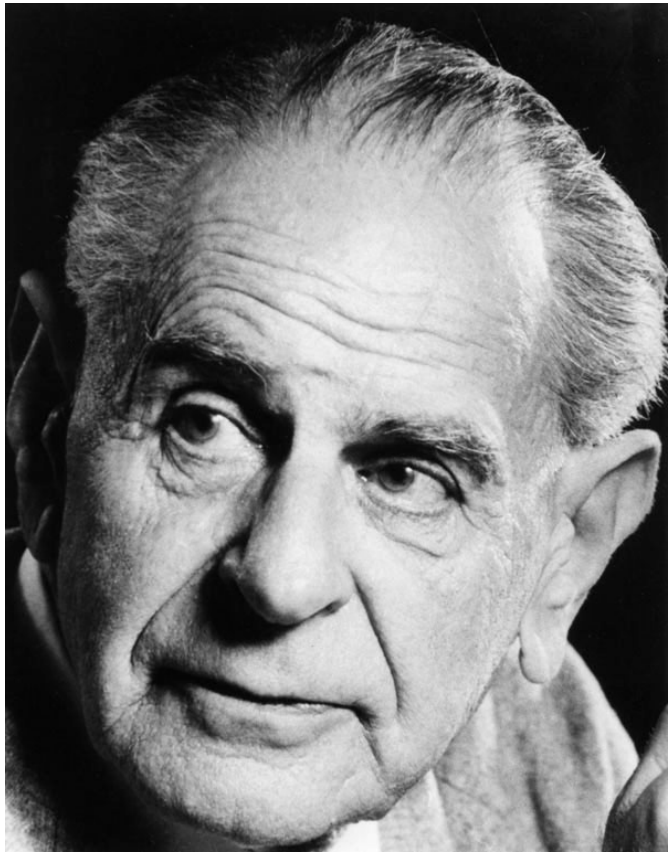


Overview

1. Philosophy, Bert Baumgaertner
2. Theory, Erkan Buzbas
3. Agent Based Model, L. Gustavo Nardin

Philosophy, Bert Baumgaertner

Karl Popper (1902-1994)



Falsifiability; inductive skeptism (confirmation is a myth); deductive logic; logical empiricism; foundationalism.

Foundationalism, Holism, and Division of Labor

- Contra Popper, confirmation *ought* to play a role.
- Holism: no hypothesis can be tested in isolation, all tests are done against a backdrop of assumptions and background knowledge (K). There is no "view from nowhere" and inferences are conditioned on K.
- Science is a collective enterprise. There is a cognitive division of labor and social structure matters.

A brief flavor of the history. . .

Robert Boyle (1627-1691)

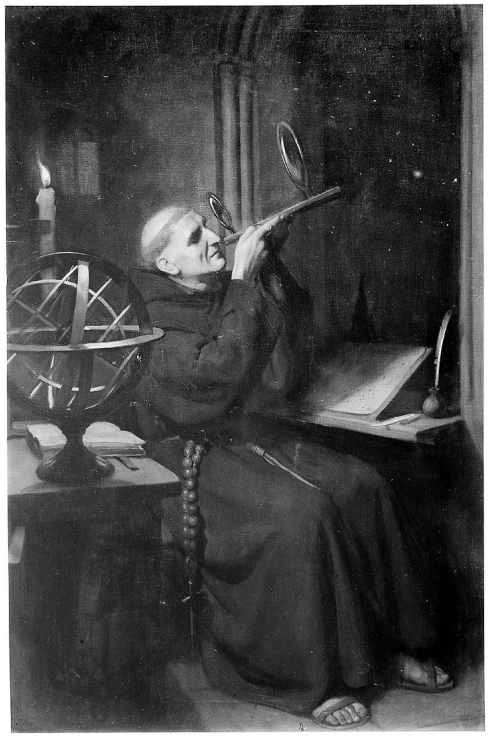
Science is a way to resolve disputes by bringing relevant questions into contact with an experimental apparatus. E.g., Aristotlean physics and vacuums.

Systematic experiment is *institutionally organized* (the Boyle circle established the Royal Society).

Francis Bacon (1561-1626)

- Inspired Boyle with idea of cooperative research institution.
- Science by "induction": gather *lots* of particular facts and observations, then generalize and make hypotheses.
- Said process requires: experimentation and *systematic reporting of results*

Roger Bacon (1214/20-1292)



- Aim: develop method of science, like logic tests for validity; certification of knowledge (halos are visual errors).
- Role of experiment: confirm, refute, or challenge theoretical claims. He repeated experiments himself.

And more…

- Avicenna (980–1037) Persian physician and philosopher; principle of replication in medical experiments.
- Alhazen (965–1040) Theoretical physicist and polymath; science is a *repeating* cycle of observations, hypothesis, experimentation, and independent verification.
- Al-Biruni (973–1048) Islamic scholar and polymath; experiments should be repeated to mitigate human and instrumentation error.
- Athenaeus of Naucratis (late 2nd to early 3rd century) Wrote of Egyptian judge ordering repeated experiments to determine if criminals survive death by venomous asps by eating lemons beforehand.
- Galen of Pergamon (c.129-200BC) Roman physician and philosopher; repeated observations.

Reproducibility in Science

Multiple independent validations of scientific claims lend credibility to those claims. This is referred to as the *reproducibility* of research findings.

- Many disciplines fail to reproduce their major findings.
- A number of factors has been suggested to contribute to nonreproducibility problem.
- A variety of solutions have been offered.
- No underlying theory guiding these efforts.

Insights from Philosophy and Motivation

1. The issues surrounding reproducibility of results are not new to science.
2. Open, collaborative practices in scientific inquiry have been intuitively proposed as a key to solve these issues.
3. A formal theory to validate this intuition is needed for a technical discussion of reproducibility.

Our programme is to provide precise definitions for concepts involved in (1) and (2), and build a theory satisfying (3).

Theory, Erkan Buzbas

Two Theoretical References

Ioannidis (2005). Why Most Published Research Findings Are False.

McElreath and Smaldino (2015). Replication, Communication, and the Population Dynamics of Scientific Discovery.

A Model-Centric View of Scientific Inquiry

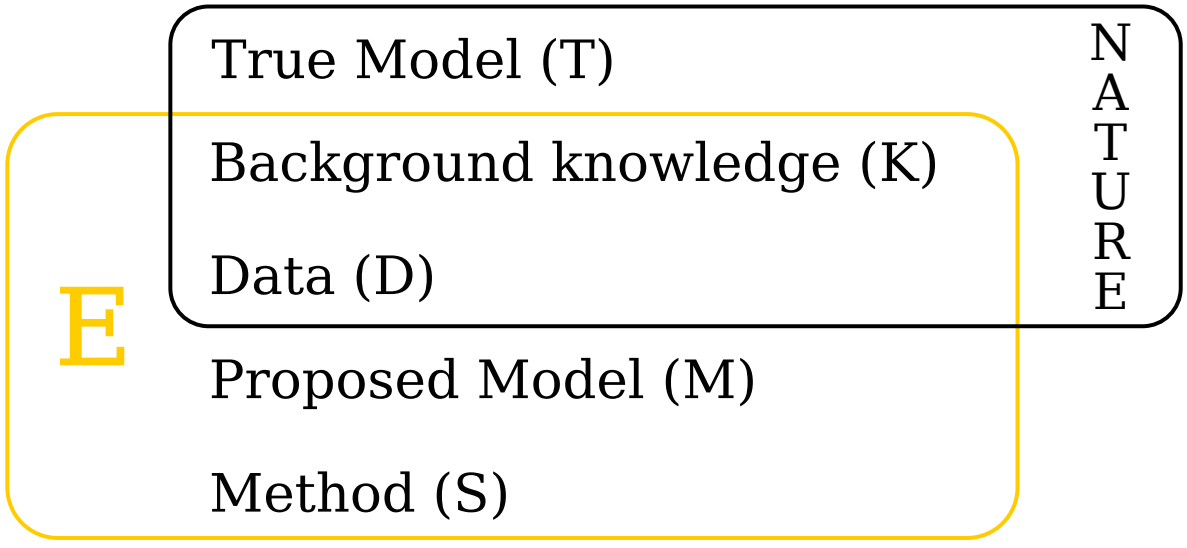
- A model-centric view of scientific inquiry:

  There exists a true model generating the data and the goal is to approximate it.

- A hypothesis-centric view of scientific inquiry:

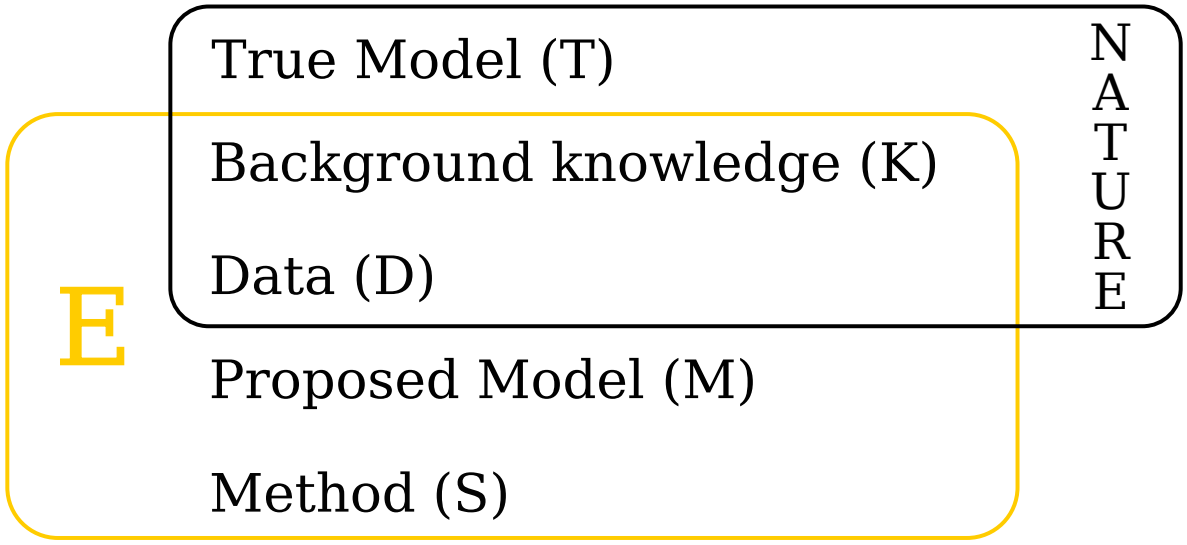
  Make claims about plausible values of parameters, assuming that a model is true.

Idealized Experiment under Model-Centric View



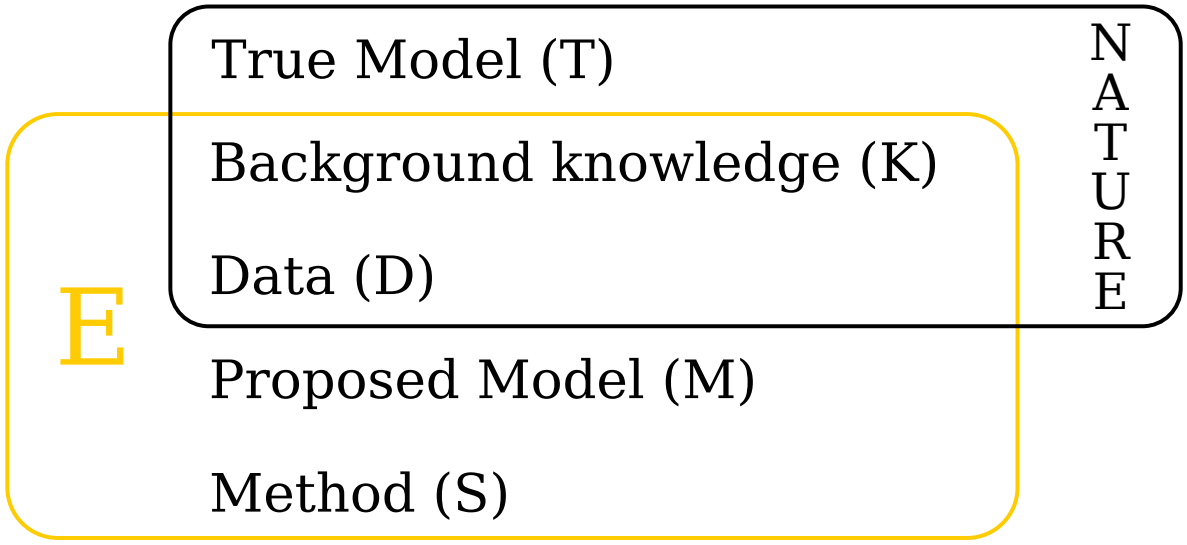
E = (K,D,M,S) is an idealized experiment.

Idealized Experiment under Model-Centric View



E = (K,D,M,S) is an idealized experiment.

Method under Model-Centric View



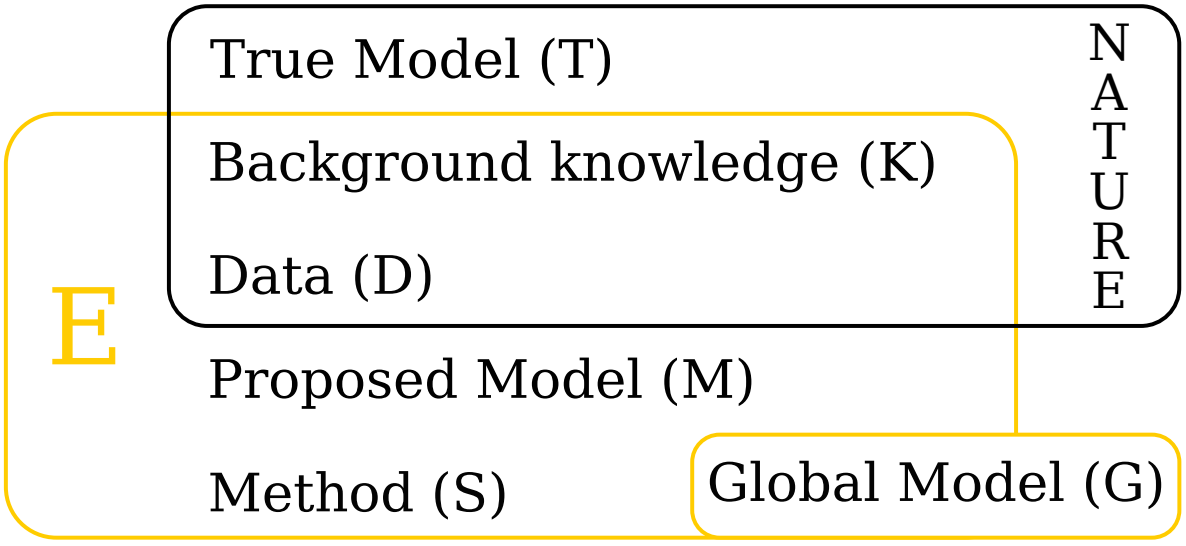
Method measures model performance:

1. Rewards model fit

$$P(M|D,K) > P(M|K)$$

2. Penalizes model complexity, everything else being equal

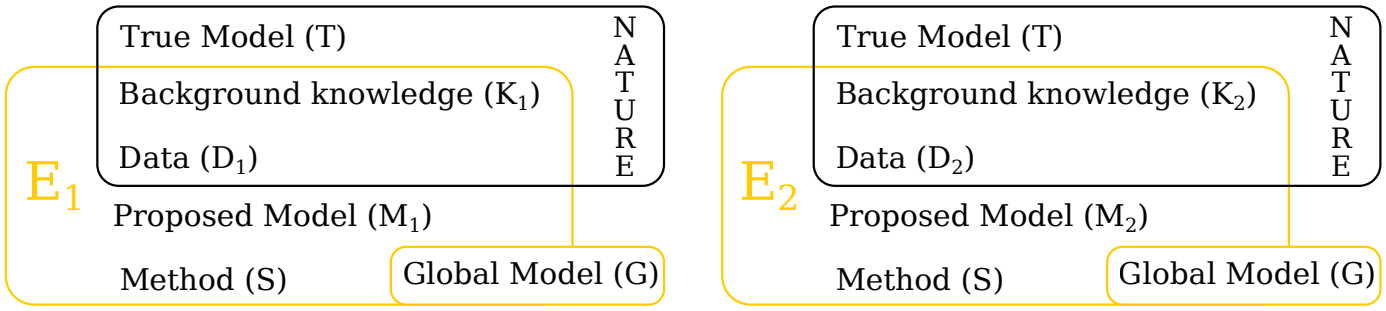
Comparison of Experiments under Model-Centric View



In comparisons of model performances,
a useful construct is a current "best" model,
which we call a Global Model (G).

Method then assigns a score S(M,G)

Reproducibility under Model-Centric View

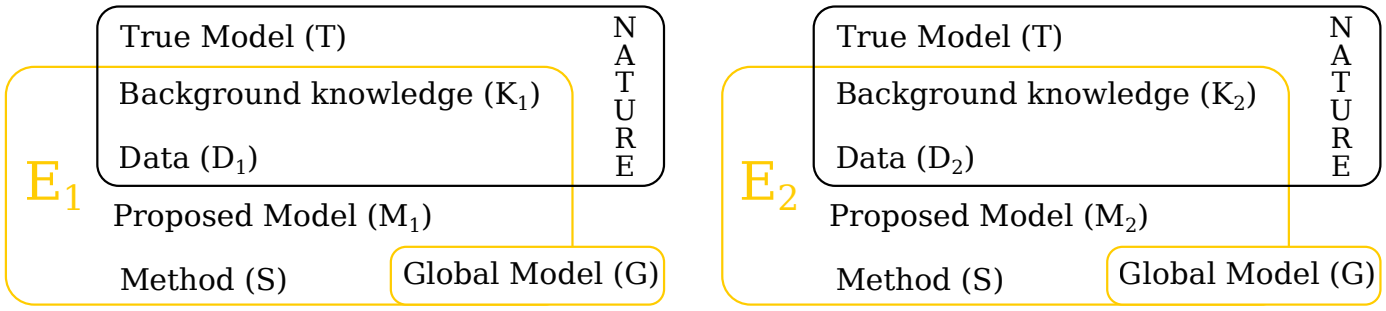


The result of E_1 is reproduced by E_2 only if:

1. $S(M_1,G)$ and $S(M_2,G)$ agree
2. E_2 knows (1)

(2) is true only if K_2 includes the information given in (1). This implies that:
Open Science is a logical necessity of reproducibility of scientific results.

Reproducibility under Model-Centric View



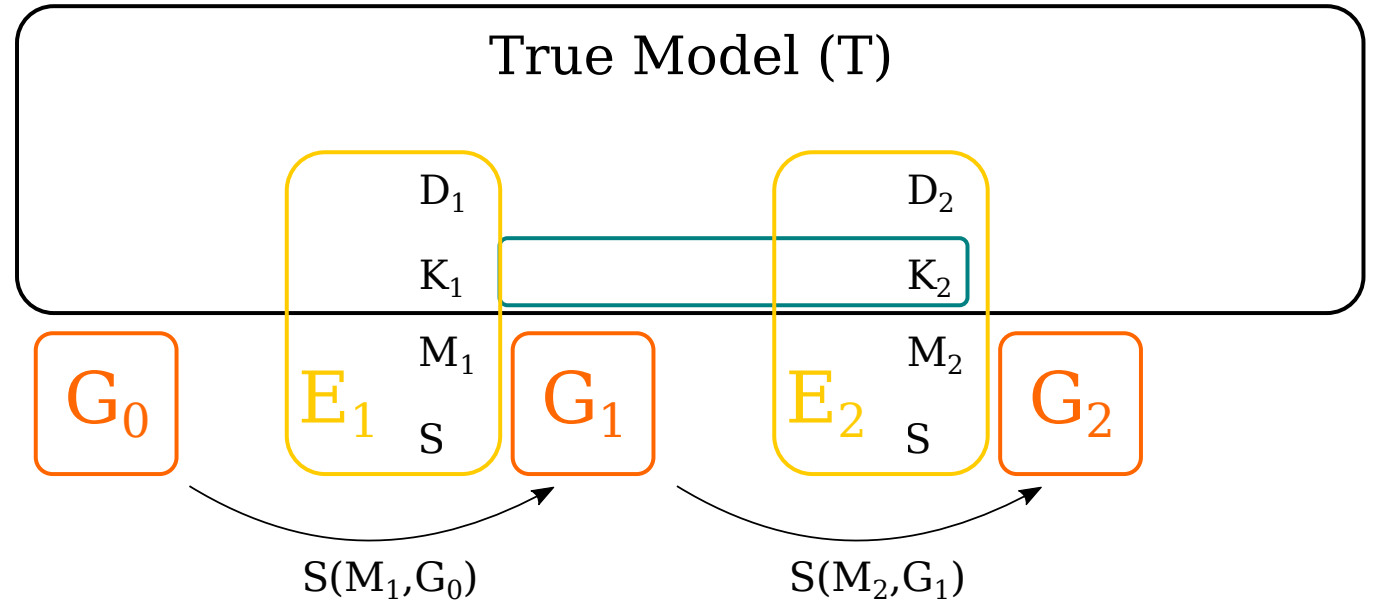
The result of E_1 is reproduced by E_2 only if:

1. $S(M_1,G)$ and $S(M_2,G)$ agree
2. E_2 knows (1)

(2) is true only if K_2 includes the information given in (1). This implies that:

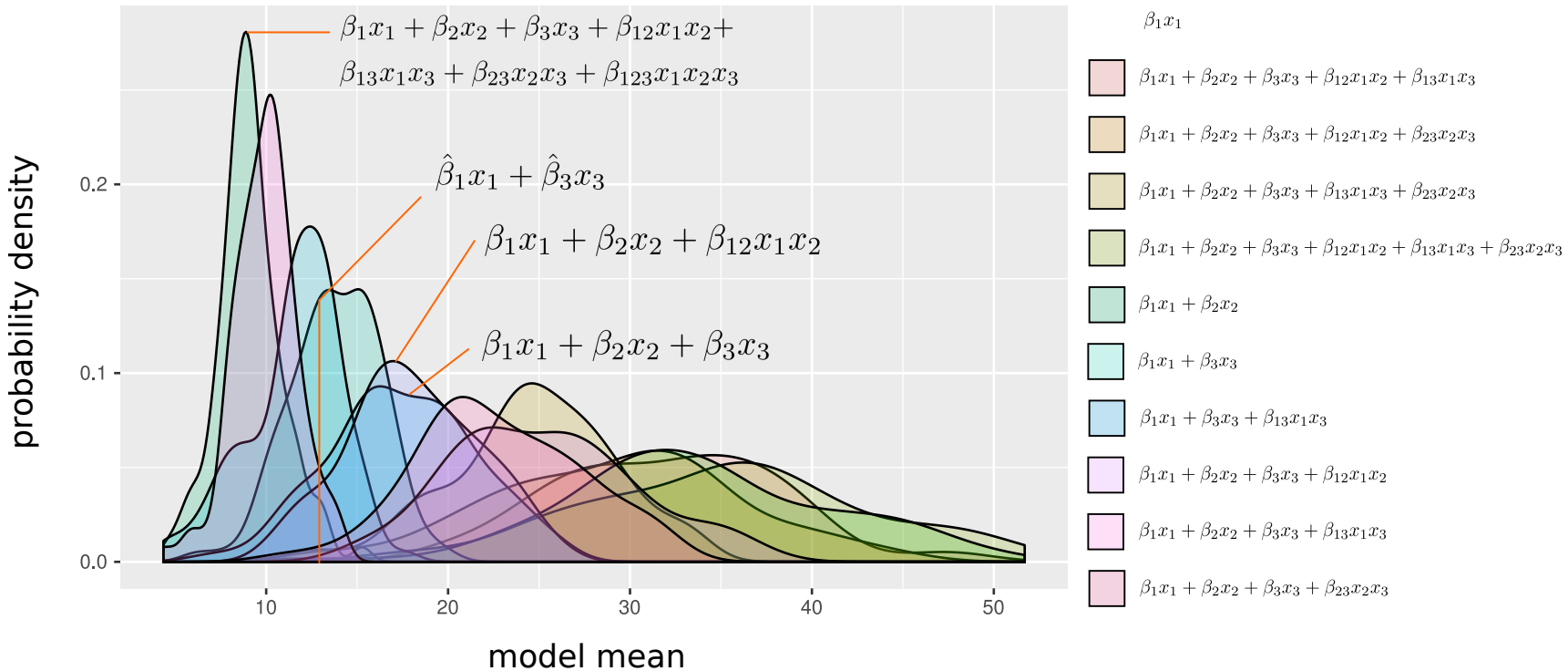
Open Science is a logical necessity of reproducibility of scientific results.

From Scientific Inquiry to Scientific Process



Global model is updated in a sequence of experiments. Later experiments have some knowledge of previous experiments. Experiments are of the same type.

A Universe of Scientific Models



Scientific Process with Multiple Strategies

Given G_t, at time t :

1. K_t uses
   - $\mathscr{M} = \{M_1, M_2, ..., M_L\}$
   - $\mathscr{R} = \{R_1, R_2, ..., R_J\}$

   to probabilistically choose a research strategy R_j and propose a model M_t
2. D_t is generated under T
3. $S(M_t, G_t)$ determines G_{t+1}
4. Convergence of the scientific community to T is monitored

Theoretical Model Probabilities

The system defines a Markov chain transitioning between global models. Transition probabilities are given by:

$$
\begin{aligned}
P(M|G_t) &= \sum_{\mathcal{R}} P(M|R, G_t)P(R) \\
&= P(M \; improves \; G_t)P(propose \; M|R, G_t)P(R) \\
&= P(S(M, G_t) > 1)\left(\frac{1}{\#\{M_R\}}\right)P(R)
\end{aligned}
$$

Theoretical Model Probabilities

Assigning positive probabilities implies the existence of a stationary distribution for this Markov chain:

In the long term, the probability of each model in the model space to become the global converges to a constant independent of the current global model.

Agent Based Model, L. Gustavo Nardin

Agent Based Model

- ABM is a *computational technique* that enables simulating the *actions* and *interactions* of agents.
- Agents are *entities* that behave on the basis of *behavioral mechanisms* and *representations*.
- ABM may be used to evaluate how different *initial conditions* and *behavioral mechanisms* affect the system outcomes. Thereby they enable
  - users to gain a deeper understand of the system's causal relationships and
  - policy-makers to analyze the effects of different policies *in silico* before intervening on the real system.

Objective

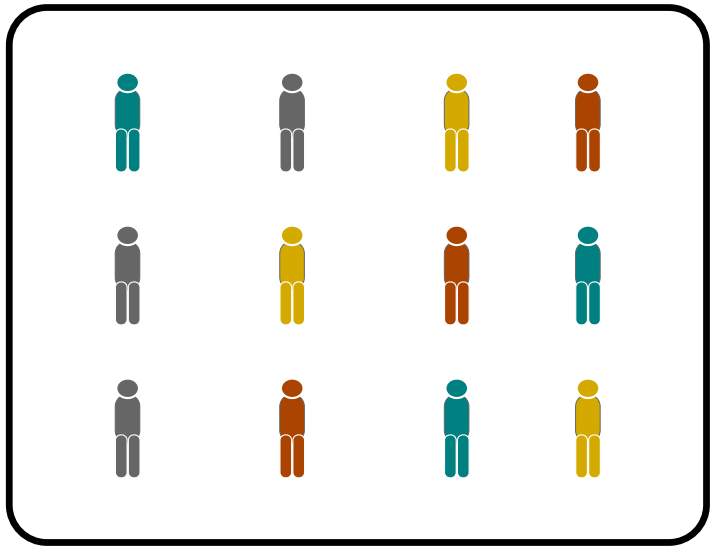
Develop an ABM representing the scientific discovery incorporating *behavioral*, *institutional*, and *methodological* aspects to inform evidence-based policies for addressing the problem of nonreproducibility.

Objective

Develop an ABM representing the scientific discovery incorporating *behavioral*~~*, institutional,* and *methodological*~~ aspects to inform evidence-based policies for addressing the problem of nonreproducibility.

ABM Description

- Set of *heterogeneous scientist agents*



ABM Description

- Set of *heterogeneous scientist agents*
- that act using *different strategies* for doing their research

ABM Description

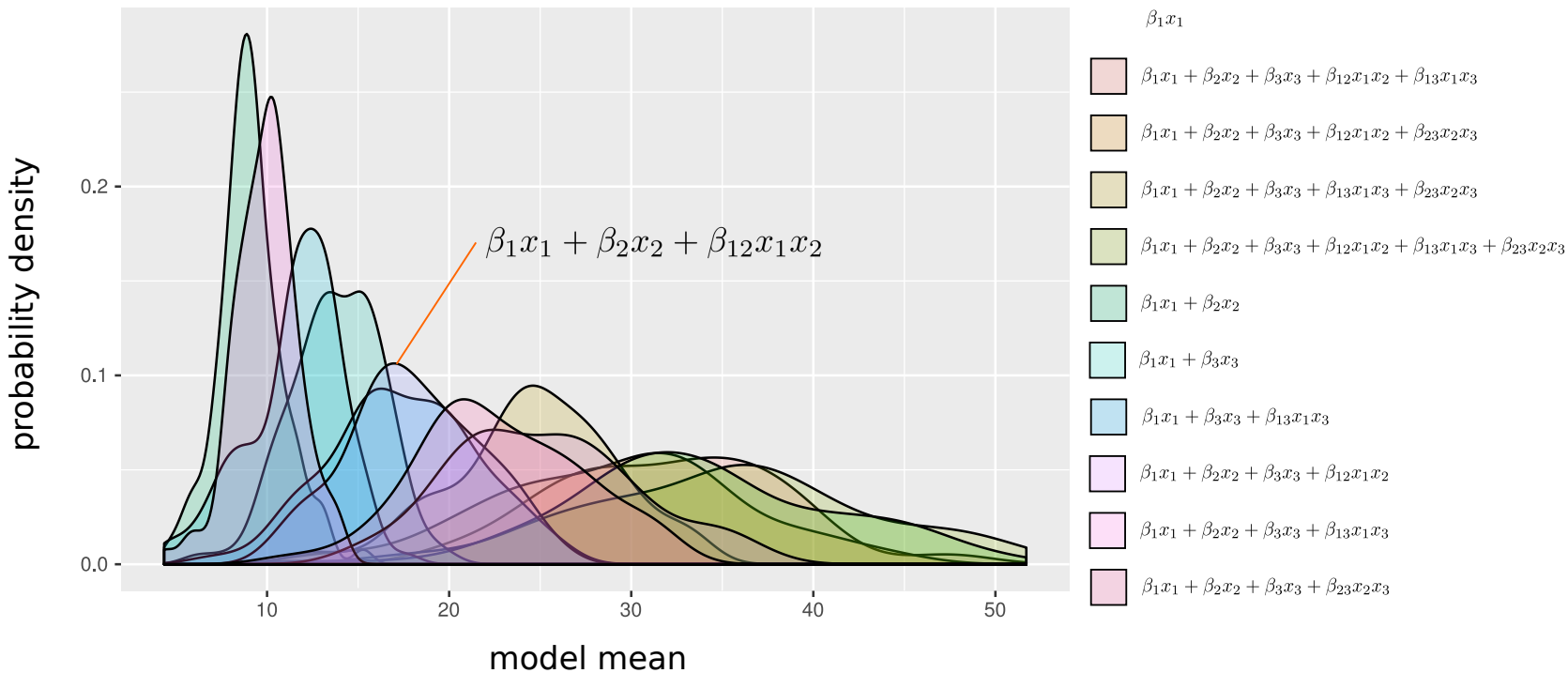
- Set of *heterogeneous scientist agents*
- that act using *different strategies* for doing their research
- aiming to *converge on the model that represents the phenomenon under investigation*, the True Model.

Research Questions

1. Does the scientific community converge on the True Model?
2. Does the True Model complexity affects the reproducibility?
3. Does scientists behavior affects the reproducibility?

ABM Structure

- *True model:* $\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$
  Model under which the investigated phenomenon generates the data.

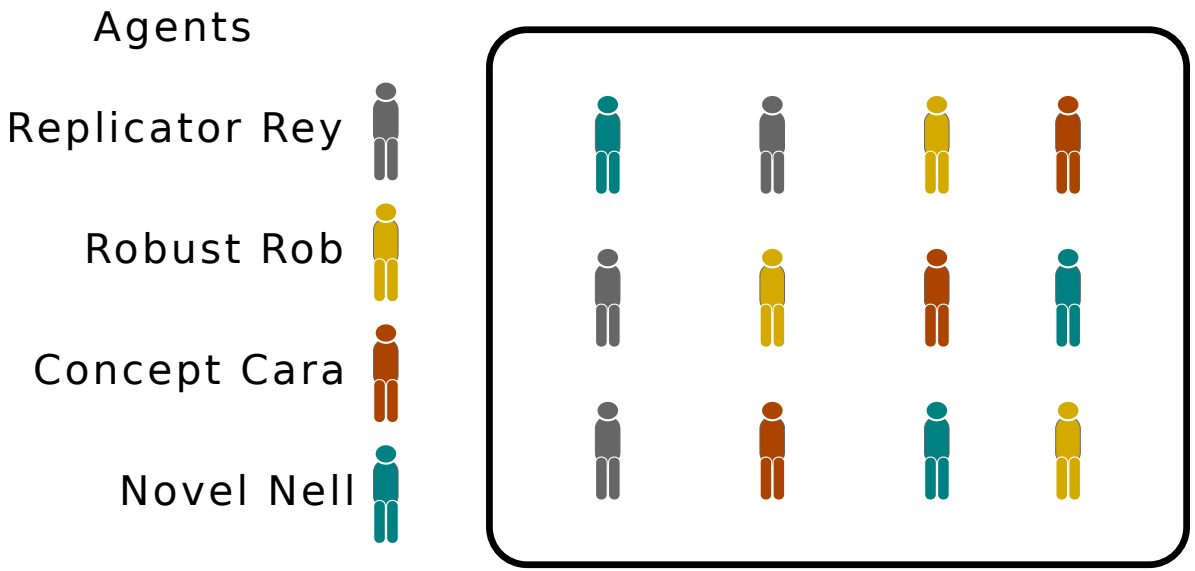


ABM Structure

- *True model:* $\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$
  Model under which the concerned natural phenomenon generates the data.
- *Global model:* $\beta_1 x_1$
  Model that the scientific community agrees on as generating the data at certain moment in time.

ABM Structure

- *True model:* $\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$
  Model under which the concerned natural phenomenon generates the data.
- *Global model:* $\beta_1 x_1$
  Model that the scientific community agrees on as generating the data at certain moment in time.
- *Scientific Community*



Scientist Agents

- **Rey** (the Replicator) conducts an *exact replication* of the study that immediately precedes her in the population.

Scientist Agents

- **Rey** (the Replicator) conducts an *exact replication* of the study that immediately precedes her in the population.
- **Rob** (the Robust) proposes a new model from the subset of models that are *one term away* from the Global Model.

Scientist Agents

- **Rey** (the Replicator) conducts an *exact replication* of the study that immediately precedes her in the population.
- **Rob** (the Robust) proposes a new model from the subset of models that are *one term away* from the Global Model.
- **Cara** (the Concept) proposes a new model by *adding a new interaction* to the Global Model.

Scientist Agents

- **Rey** (the Replicator) conducts an *exact replication* of the study that immediately precedes her in the population.
- **Rob** (the Robust) proposes a new model from the subset of models that are *one term away* from the Global Model.
- **Cara** (the Concept) proposes a new model by *adding a new interaction* to the Global Model.
- **Nell** (the Novel) does not build off of the literature but comes up with her original ideas, *choosing a random model* from the Universe of Models.

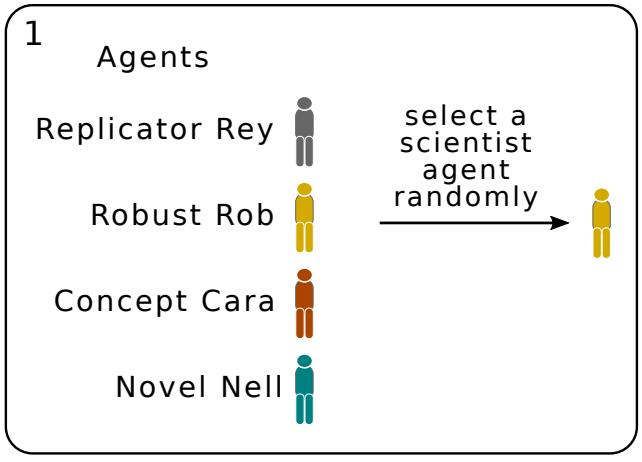
ABM Dynamics: Initialization

Given a set of input parameters:

- generates the Scientific Community (*Prop. Agent Strategies)*.
- generates the Universe of Scientific Models (*Predictors*).
- defines the *deterministic part* of the True Model by randomly generating
  - a set of *predictor values* (*Sample Size* and *Correlation*), and
  - β values (*Error Variance*).

ABM Dynamics (1)

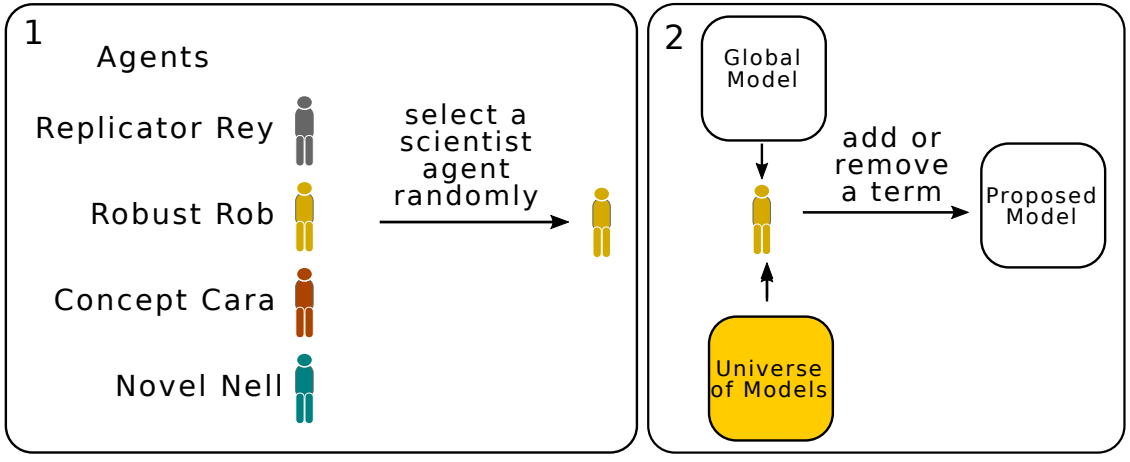
- Selects a Scientist Agent randomly



- True Model: $\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$
- Global Model: $\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2$

ABM Dynamics (2)

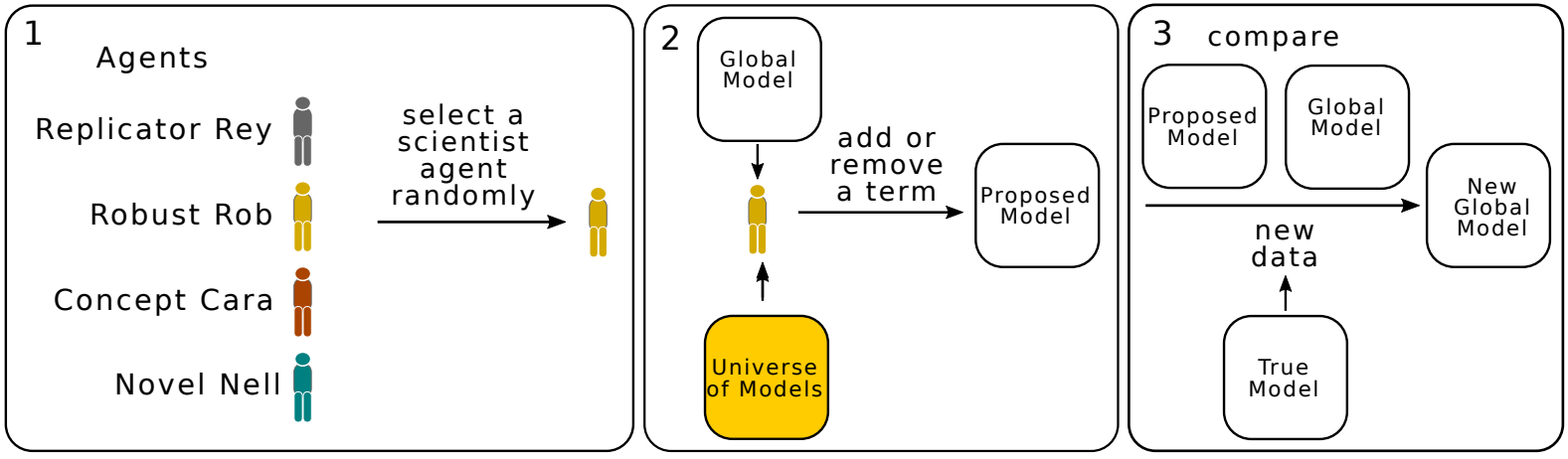
- The Scientist Agent based on the current Global Model and its own strategy proposes a new model



- True Model: $\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$
- Global Model: $\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2$
- Proposed Model: $\beta_1 x_1 + \beta_3 x_3$

ABM Dynamics (3)

- The quality of the Proposed and the Global Model are compared for the new data generated by the True Model (e.g, *AIC*, *R-Square*, *T Statistics*)



- True Model: $\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$
- Global Model: $\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2$
- Proposed Model: $\beta_1 x_1 + \beta_3 x_3$
- New Global Model: $\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2$

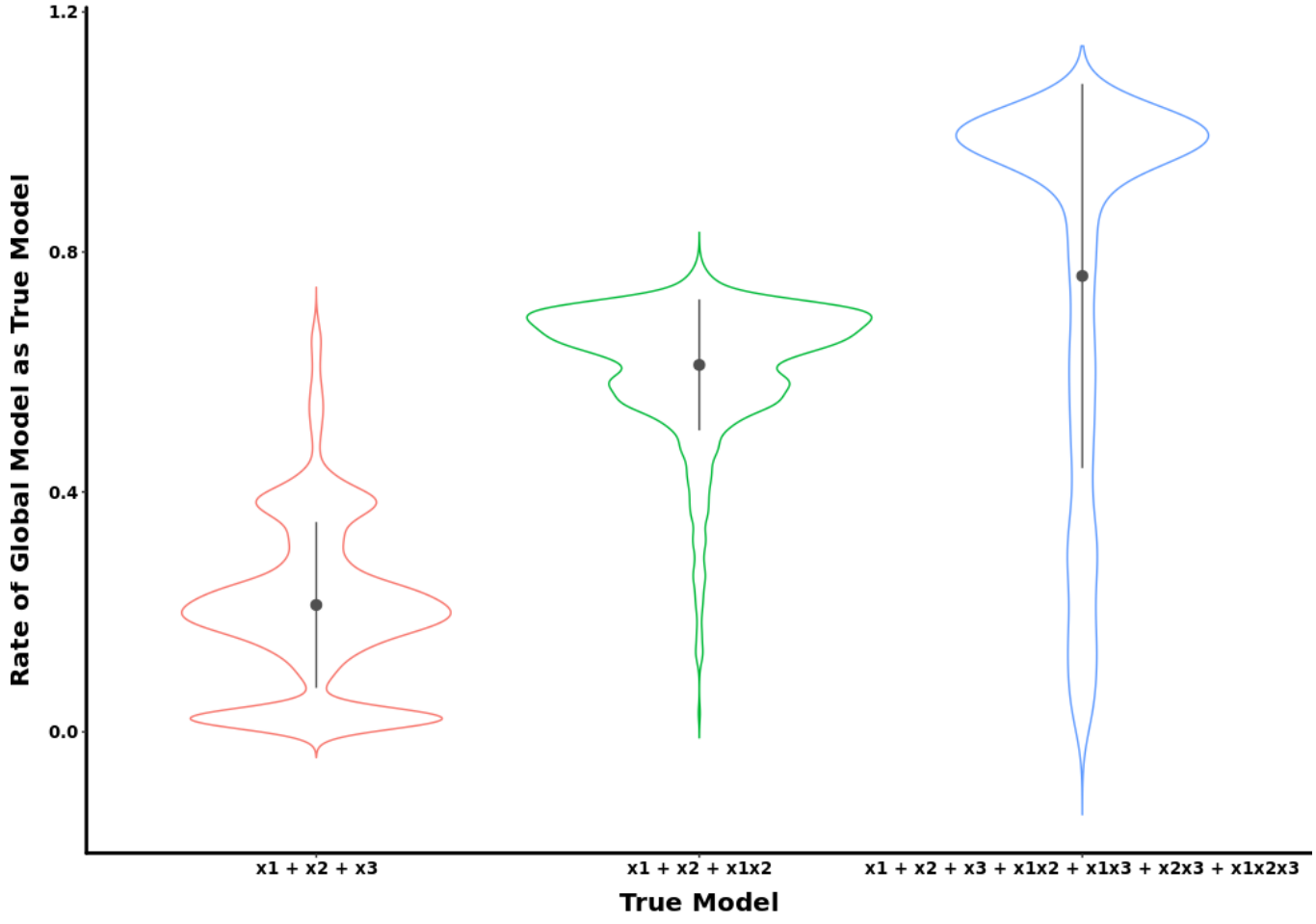
Experiment

Parameter	**Values**
Replications	100
Timesteps	10000
Universe of Models	14 (3 predictors)
Sample Size	100
Comparison Method	AIC
True Model	$\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$ $\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ $\beta_1 x_1 + \beta_1 x_2 + \beta_1 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3$
Error Variance	20% 60%
Correlation	0.2 0.8

Experiment

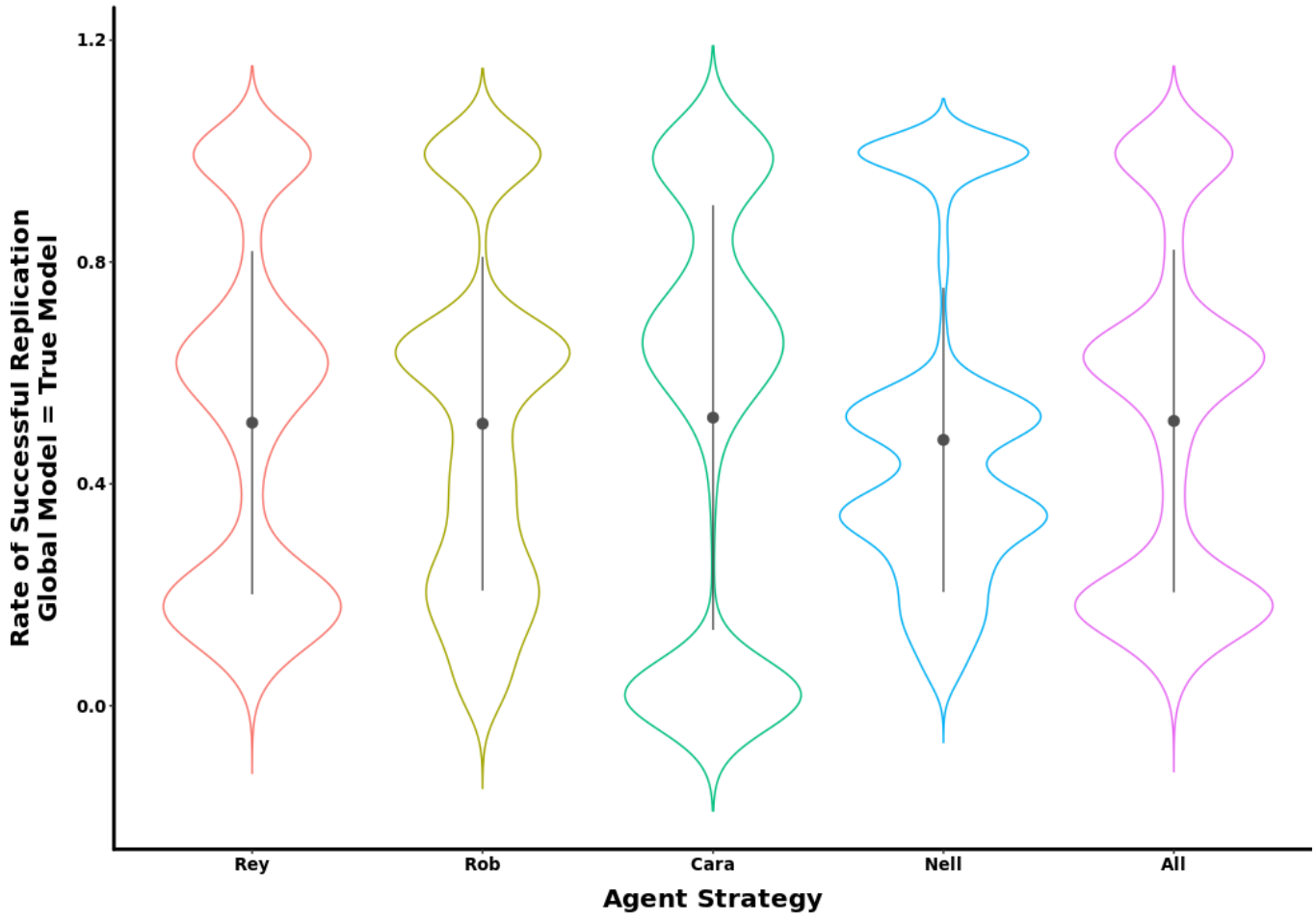
Parameter	Values
Agent Types	5 combinations ○ 4 combinations in which the Scientific Community is dominated by 85% of one Agent Strategy and 5% of the other three Agent Strategies ○ 1 combination in which the Scientific Community is composed of 25% of each Agent Strategy

Results: Main Effect of the True Model Complexity



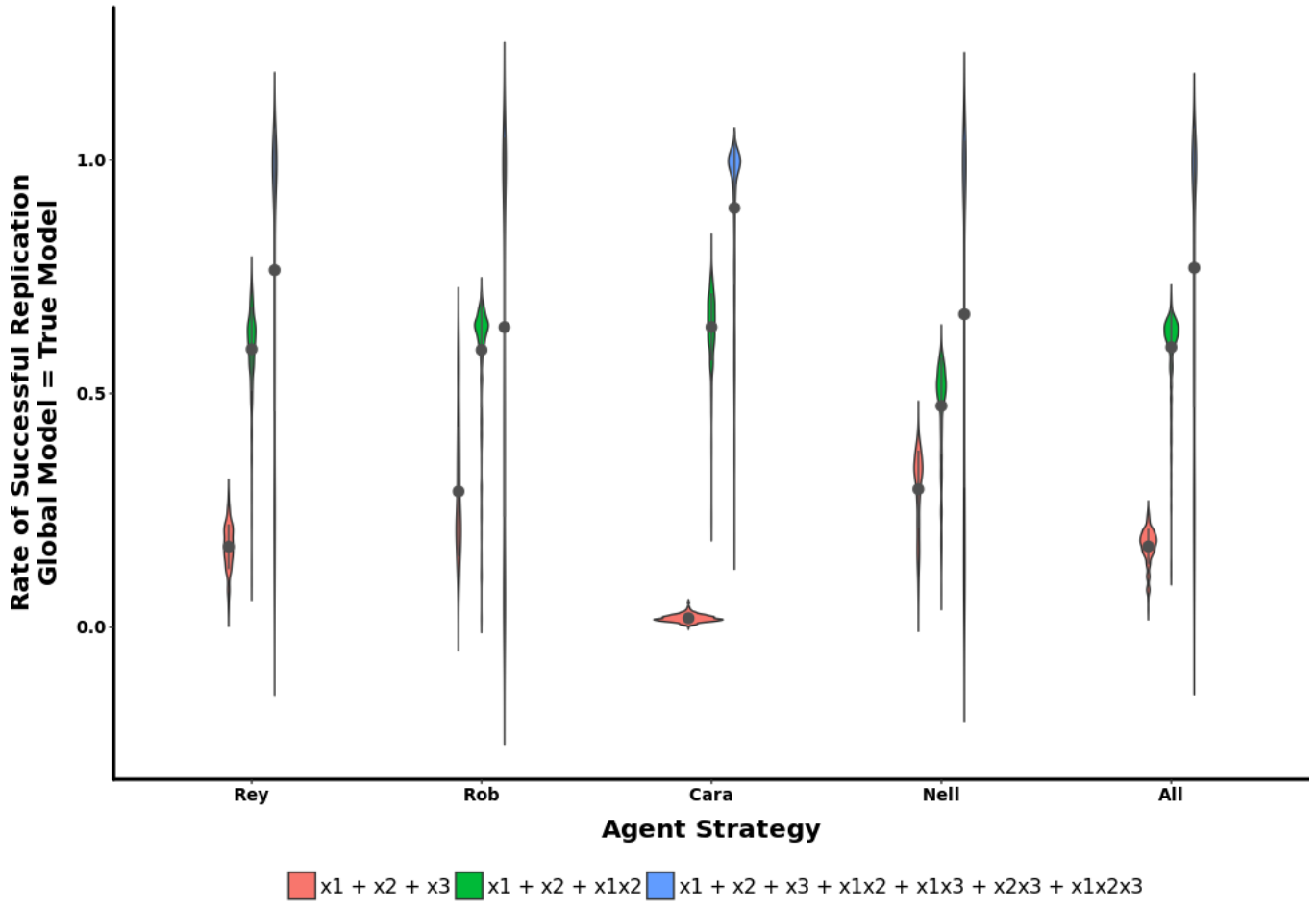
The more complex the True Model, the longer it stays as the Global Model.

Results: Main Effect of the Agent Strategy



Although they have similar mean and standard deviation, the blobs indicate there are interaction influences.

Results: True Model and Agent Strategy Interaction



Cara leads to the lowest proportion of True Model selection as Global Model under the simplest True Model because the True Model lacks interactions.

Research Questions

1. **Does the scientific community converge on the True Model?**
   Yes, although the convergence is not always stable

Research Questions

1. **Does the scientific community converge on the True Model?**
   Yes, although the convergence is not always stable
2. **Does the True Model complexity affects the reproducibility?**
   Yes, it does.
   Simple True Models tend to be reproduced less than more complex models.

Research Questions

1. **Does the scientific community converge on the True Model?**
   Yes, although the convergence is not always stable
2. **Does the True Model complexity affects the reproducibility?**
   Yes, it does.
   Simple True Models tend to be reproduced less than more complex models.
3. **Does scientists behavior affects the reproducibility?**
   The dominant Agent Strategy has a small effect on the reproducibility of True Models.

Web-based tool: CRUST

Conceptualizing Reproducibility Using Simulations and Theory