Minimizing Noise and Bias in Human and AI Decision-Making

Vasant Dhar

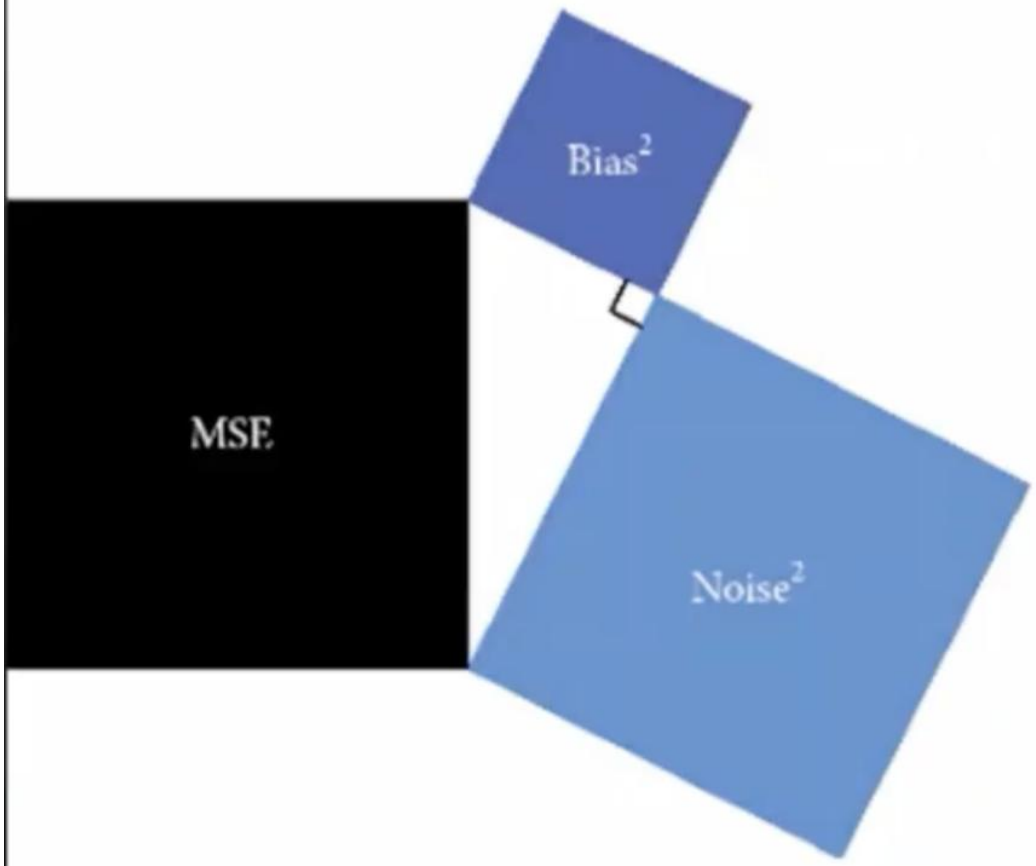
NYU Stern School of Business & Center for Data Science

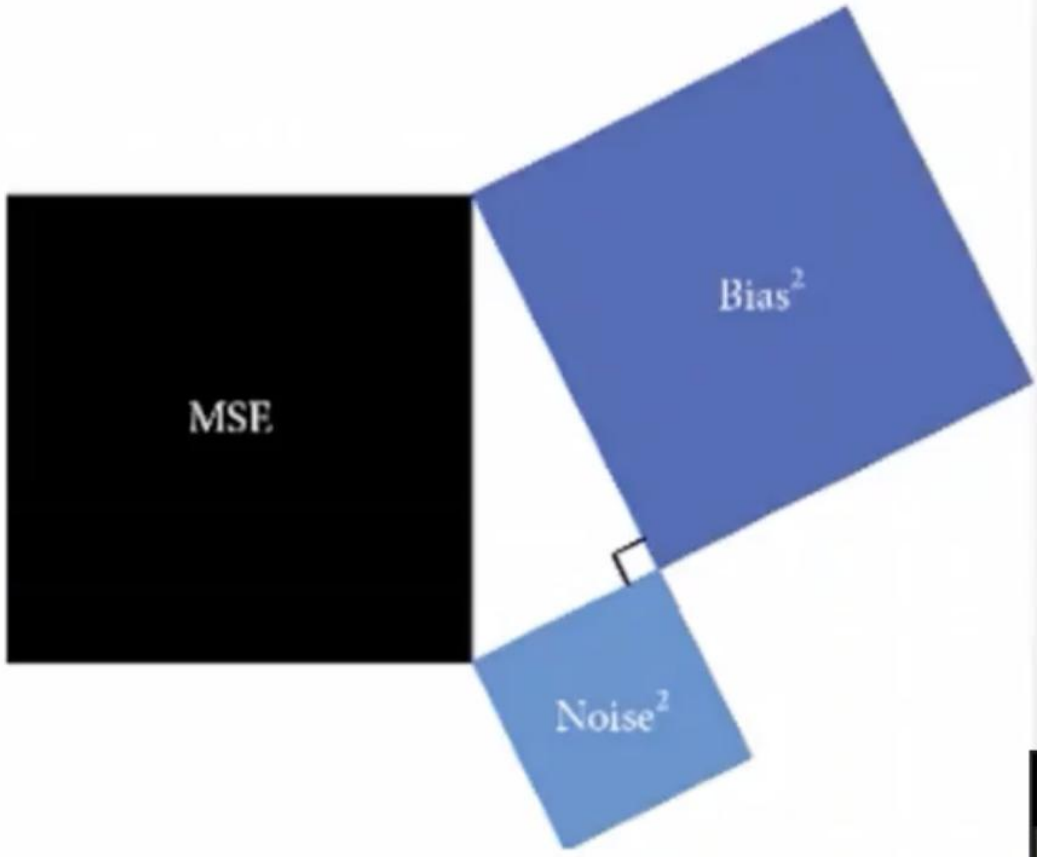
Host of the podcast "Brave New World"

Bravenewpodcast.com

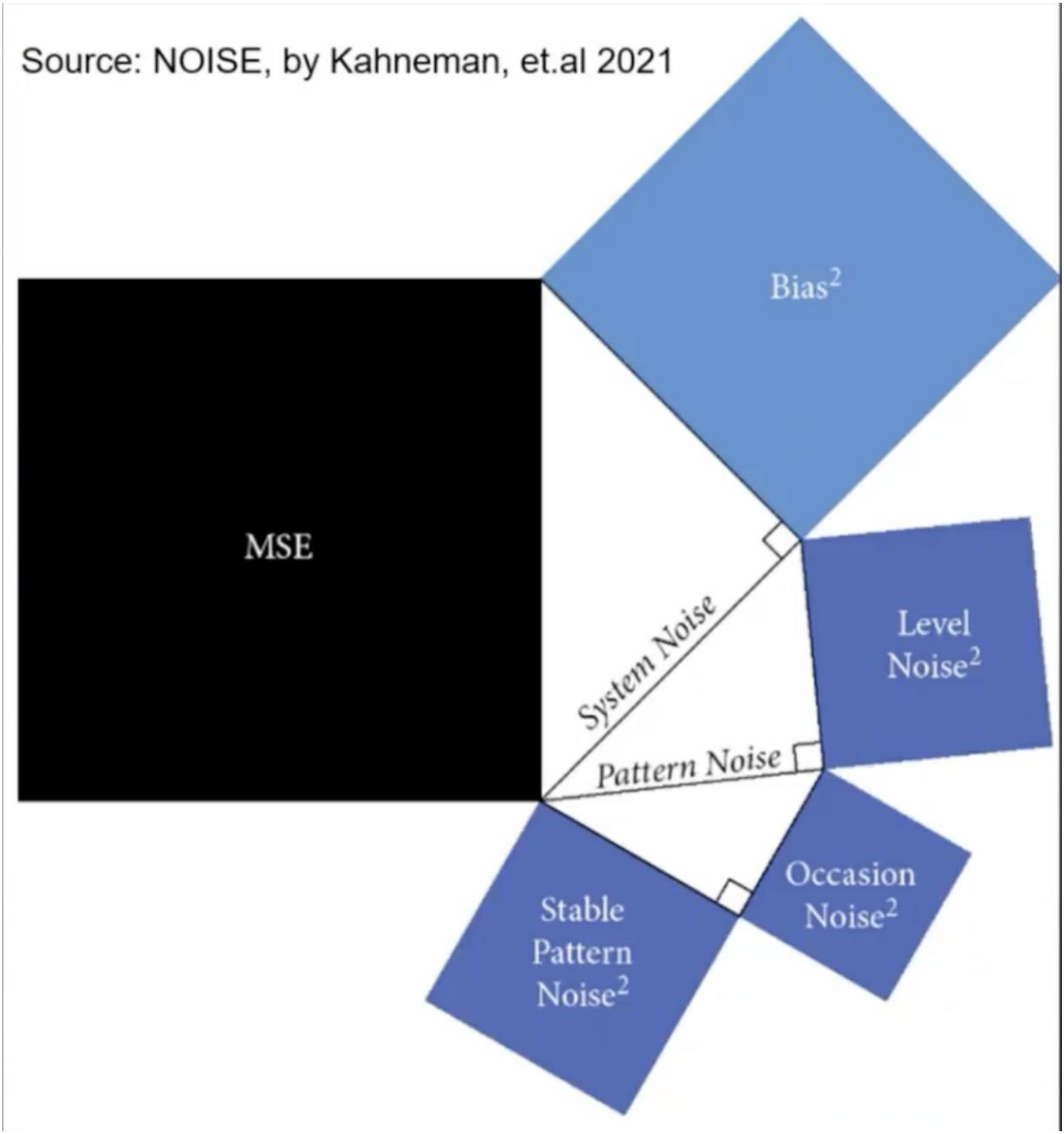
@vasantdhar

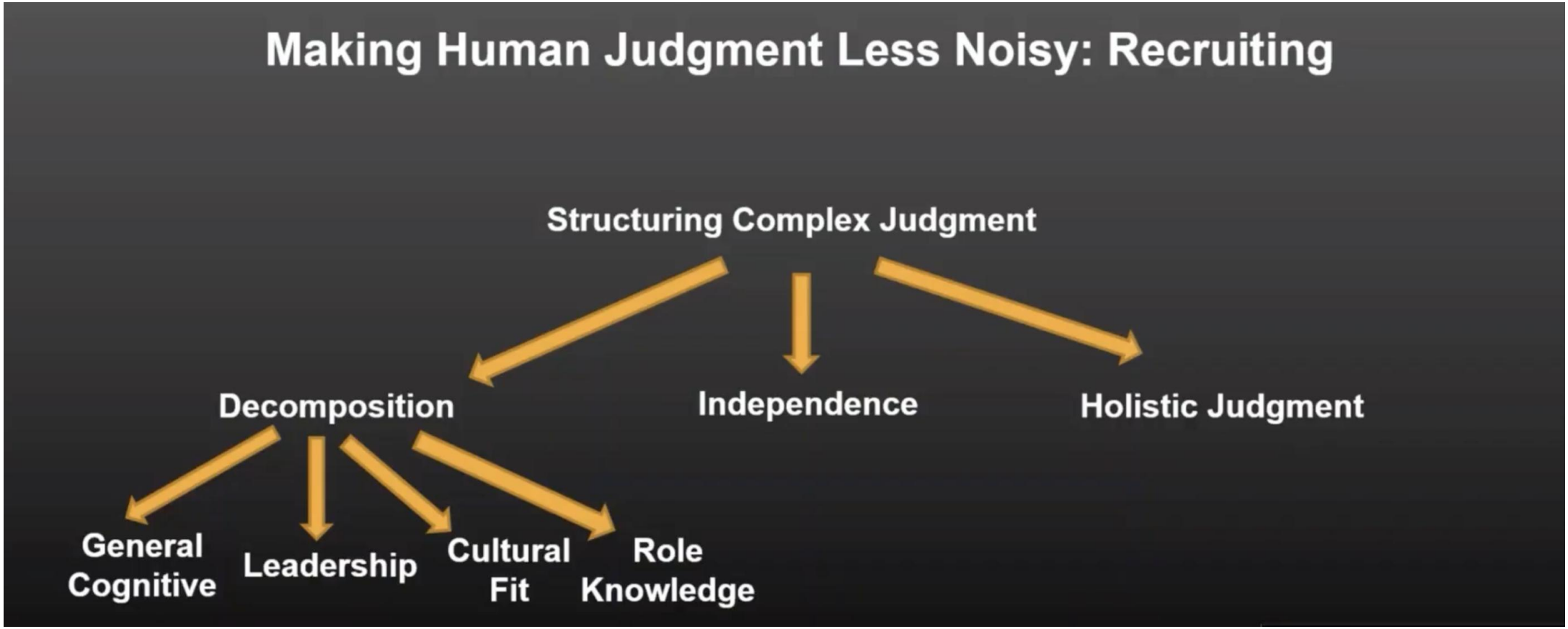
Sources of Error



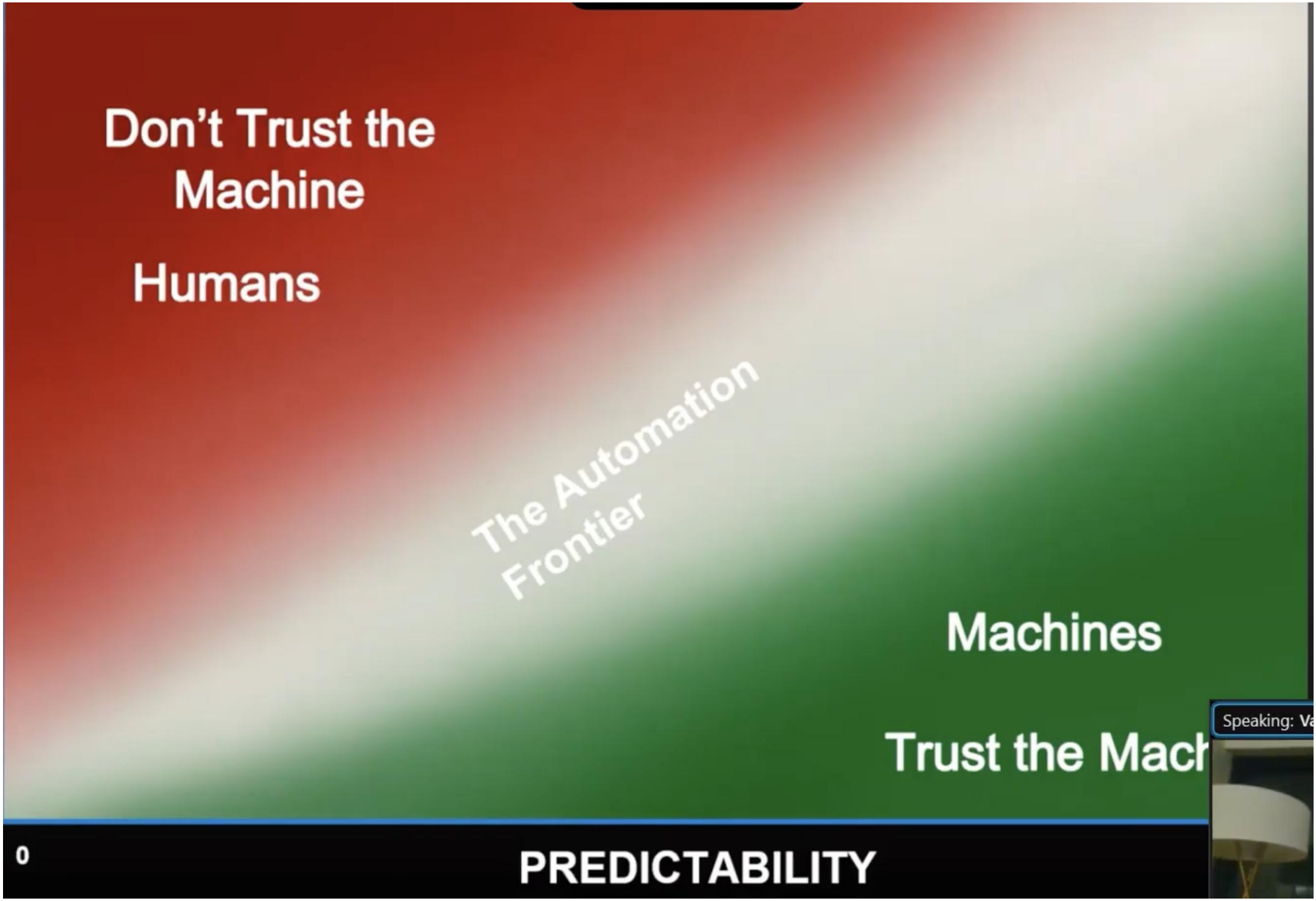


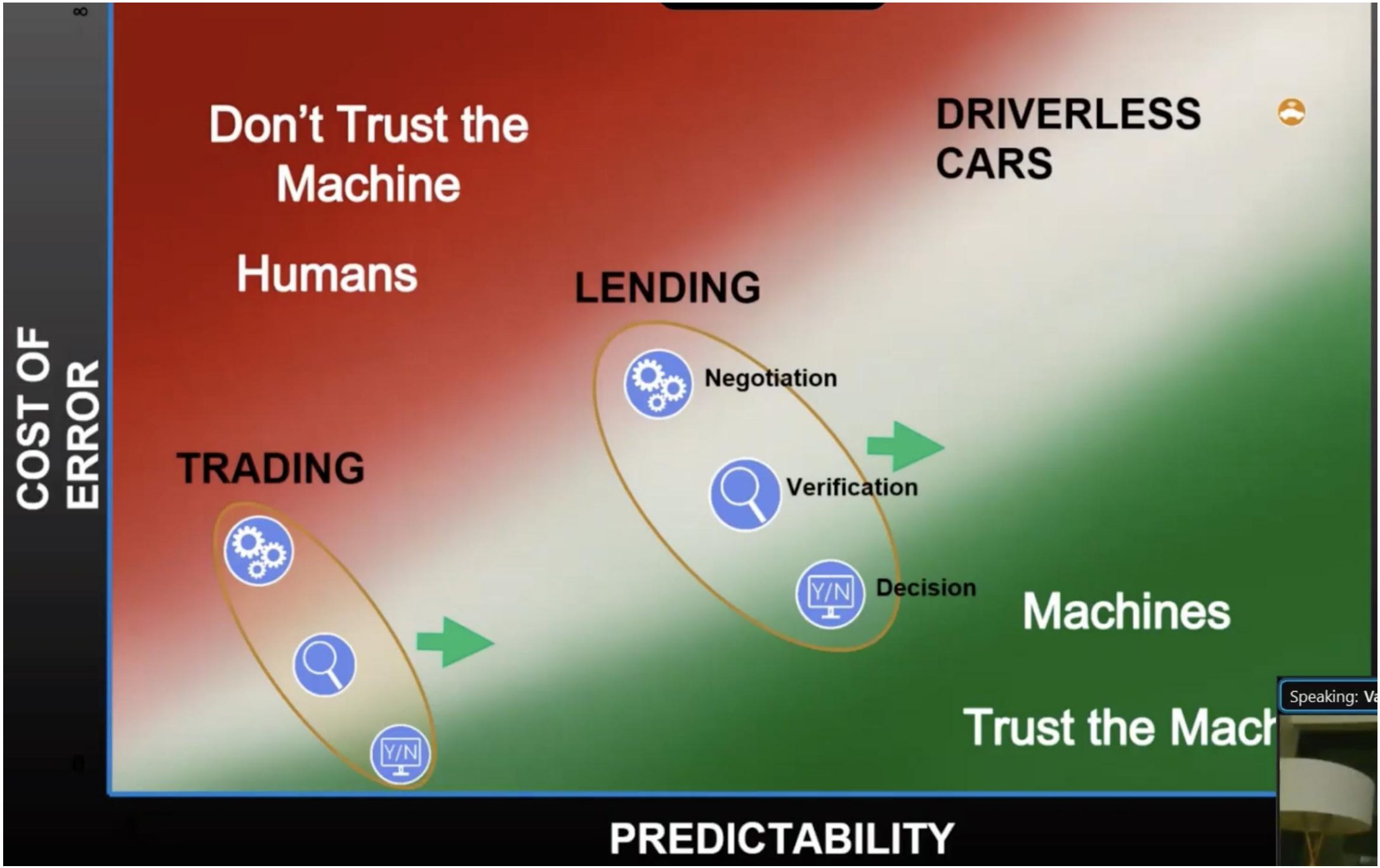
Source: NOISE, by Kahneman, et.al 2021




Making Human Judgment Less Noisy: Recruiting
Structuring Complex Judgment
Decomposition
Independence
Holistic Judgment
General Cognitive
Leadership
Cultural Fit
Role Knowledge

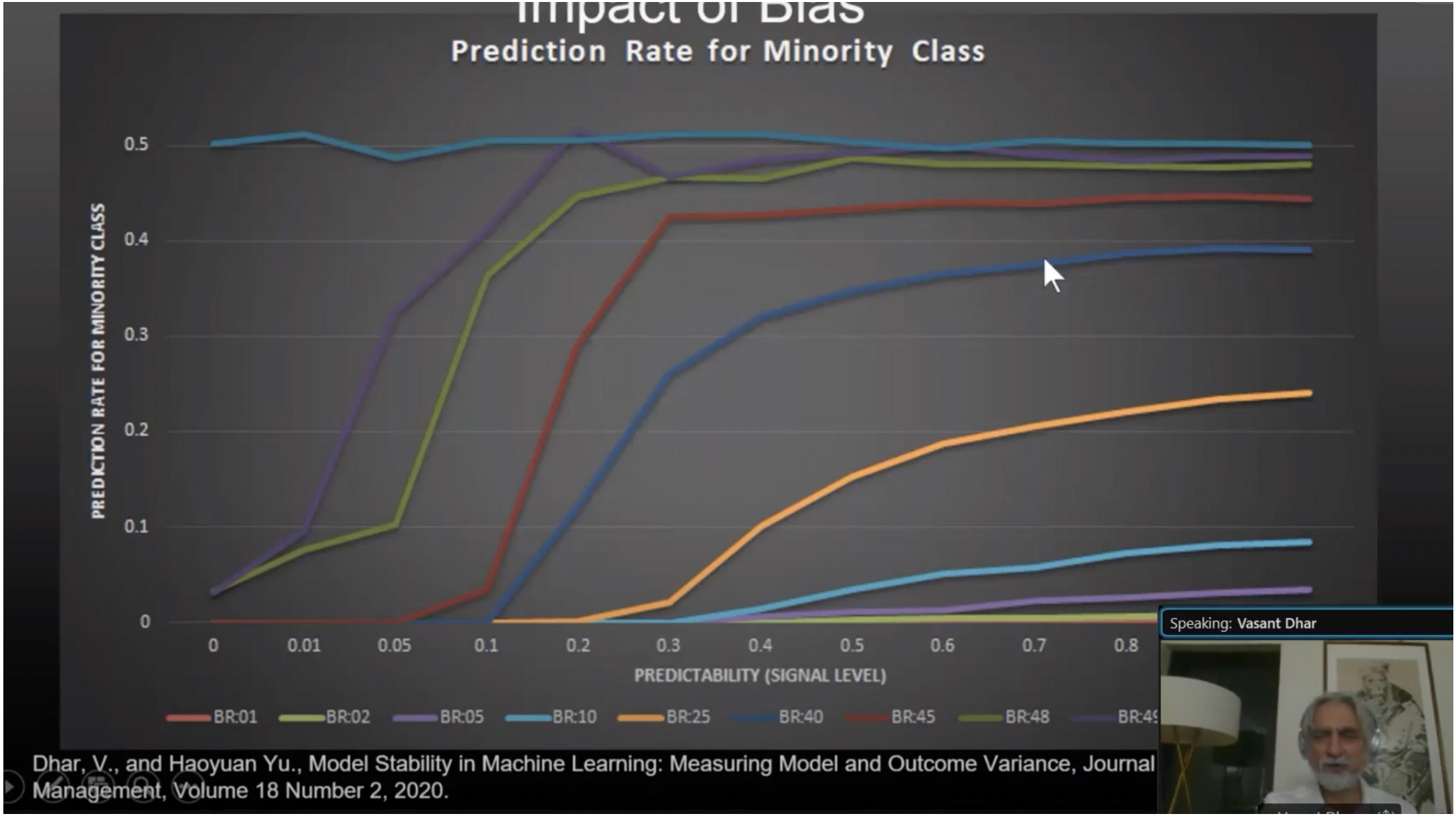
Machine-Based Decision-Making


Don’t Trust the Machine
Humans
The Automation Frontier
Machines
Trust the Mach
Speaking: Va
0
PREDICTABILITY


Don't Trust the Machine
Humans
DRIVERLESS CARS
LENDING
Negotiation
Verification
Decision
TRADING
Y/N
Y/N
Machines
Trust the Mach
COST OF ERROR
PREDICTABILITY
Speaking: Va

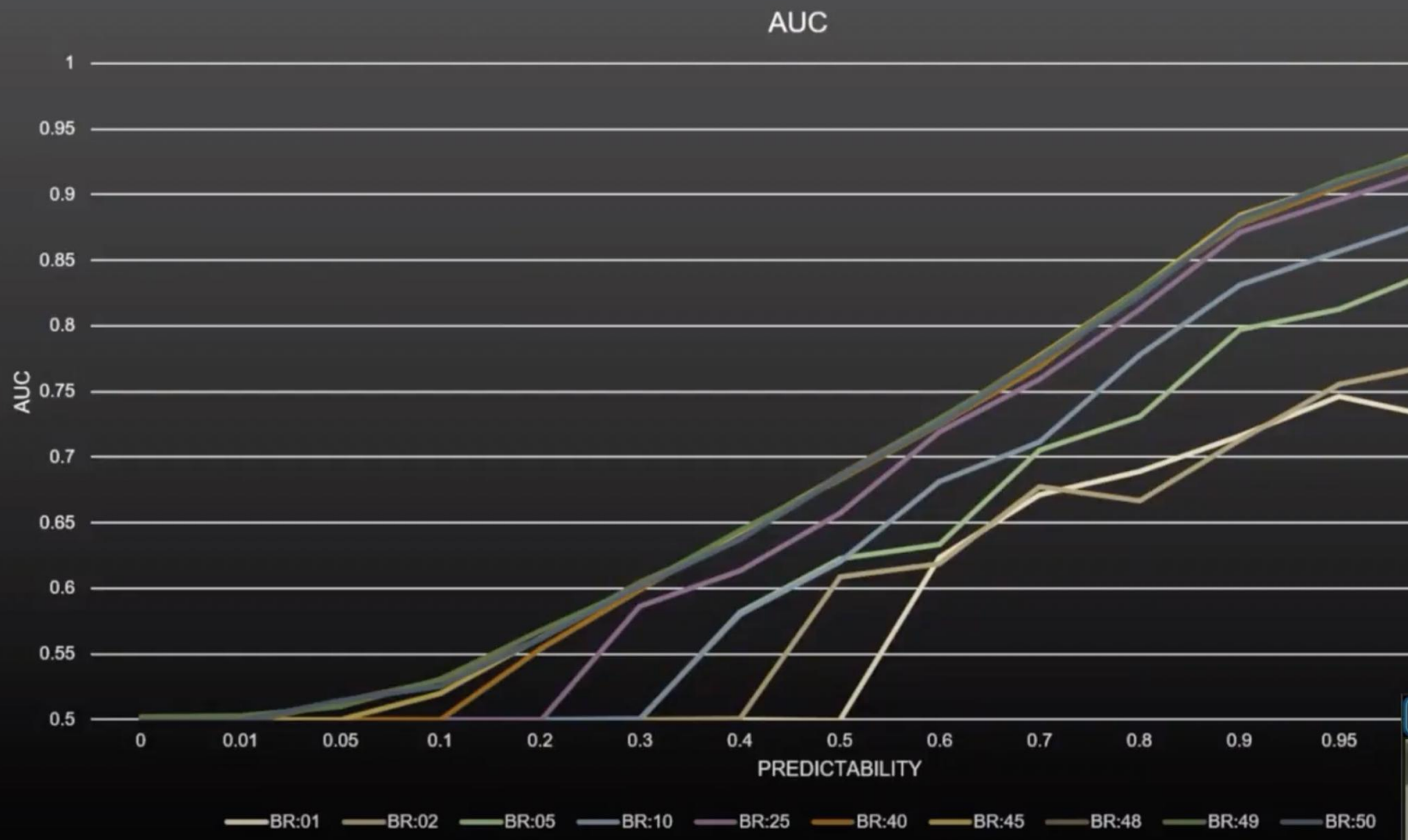
There's no Way of Getting Away!

- Noise
  - Sampling!
  - Non-stationarity!
- Bias
  - Trend
  - Observational


Impact of Bias
Prediction Rate for Minority Class
PREDICTION RATE FOR MINORITY CLASS
0.5
0.4
0.3
0.2
0.1
0
0
0.01
0.05
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
PREDICTABILITY (SIGNAL LEVEL)
BR:01
BR:02
BR:05
BR:10
BR:25
BR:40
BR:45
BR:48
Dhar, V., and Haoyuan Yu., Model Stability in Machine Learning: Measuring Model and Outcome Variance, Journal
Management, Volume 18 Number 2, 2020.
Speaking: Vasant Dhar

Viewing Vasant Dhar's screen

Impact of Bias

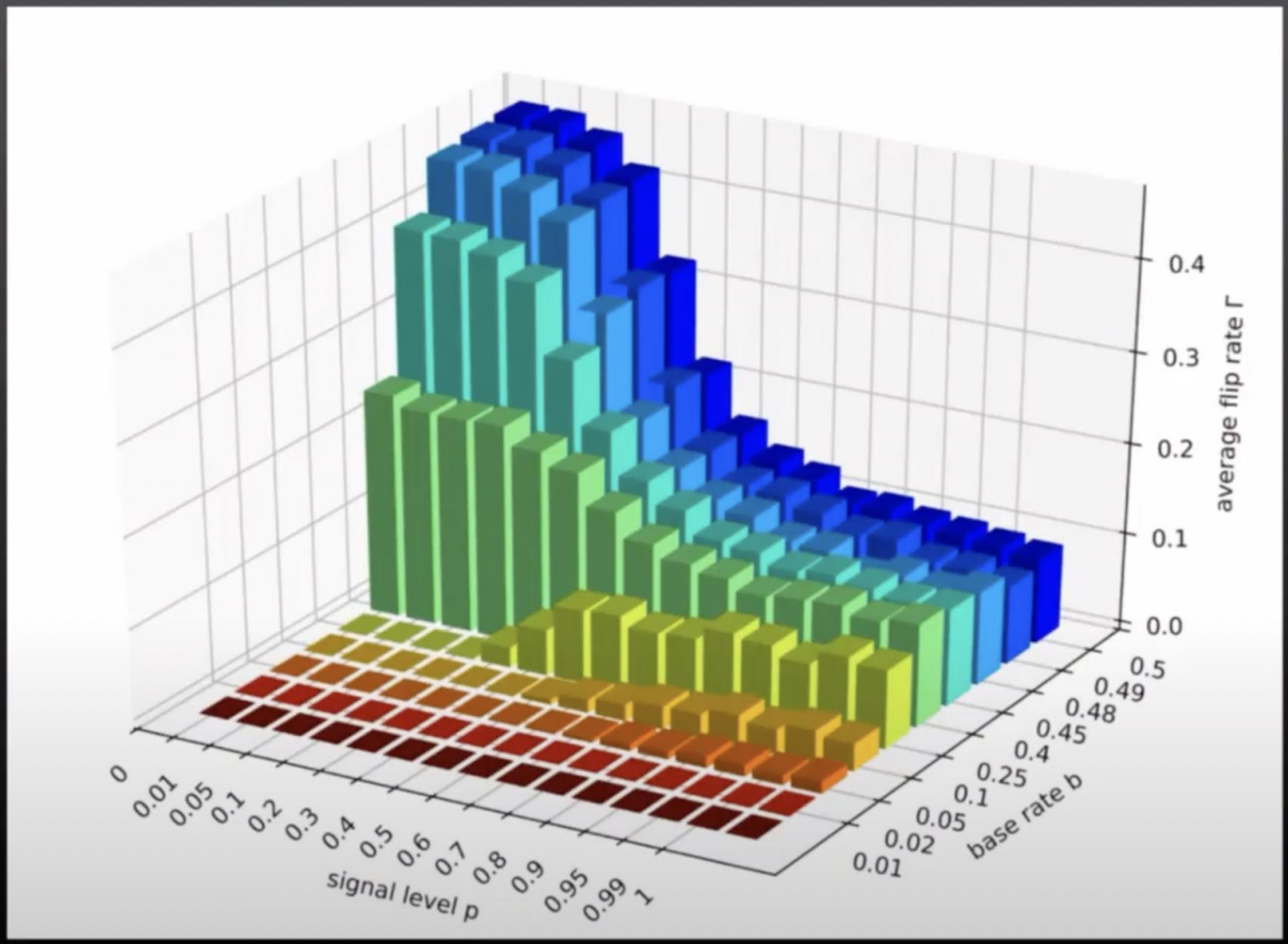


Dhar, V., and Haoyuan Yu., Model Stability in Machine Learning: Measuring Model and Outcome Variance, Journal Management, Volume 18 Number 2, 2020.

Speaking: **Vasant Dhar**

Viewing Vasant Dhar's screen

Impact of Noise: "Model Variance"



Dhar, V., and Haoyuan Yu., Model Stability in Machine Learning: Measuring Model and Outcome Variance, Journal Management, Volume 18 Number 2, 2020.

Speaking: **Vasant Dhar**

So, what do you do about it?!!!

My approach: “FastML!”

The Human-Machine Interface

- When do we need human judgment?
  - for the "edge" cases (high cost of error)

What should the human-machine interface look like?

example the Boeing 737 MAX disasters of 2018/2019

  - sensing "regime shifts"

Tetlock's work on "Superforecasters" shows that some humans have "skill" in forecasting events well into the future