# Decision Processes Colloquia

Monday, September 30, 2024

Where: JMHH F 55

When: 12:00 – 1:20 pm

## David Rand

Erwin H. Schell Professor; Professor of Marketing
Massachusetts Institute of Technology

## Durably reducing conspiracy beliefs through dialogues with AI

ABSTRACT:

Conspiracy theories are a paradigmatic example of beliefs that, once adopted, are extremely difficult to dispel. Influential psychological theories propose that conspiracy beliefs are uniquely resistant to counterevidence because they satisfy important needs and motivations. Here, we raise the possibility that previous attempts to correct conspiracy beliefs have been unsuccessful merely because they failed to deliver counterevidence that was sufficiently compelling and tailored to each believer's specific conspiracy theory (which vary dramatically from believer to believer). To evaluate this possibility, we leverage recent developments in generative artificial intelligence (AI) to deliver well-argued, person-specific debunks to a total of N = 2,190 conspiracy theory believers. Participants in our experiments provided detailed, open-ended explanations of a conspiracy theory they believed, and then engaged in a 3-round dialogue with a frontier generative AI model (GPT-4 Turbo) which was instructed to reduce each participant's belief in their conspiracy theory (or discuss a banal topic in a control condition). Across two experiments, we find robust evidence that the debunking conversation with the AI reduced belief in conspiracy theories by roughly 20%. This effect did not decay over 2 months' time, was consistently observed across a wide range of different conspiracy theories, and occurred even for participants whose conspiracy beliefs were deeply entrenched and of great importance to their identities. Furthermore, although the dialogues were focused on a single conspiracy theory, the intervention spilled over to reduce beliefs in unrelated conspiracies, indicating a general decrease in conspiratorial worldview, as well as increasing intentions to challenge others who espouse their chosen conspiracy. These findings highlight that even many people who strongly believe in seemingly fact-resistant conspiratorial beliefs can change their minds in the face of sufficient evidence.

Wharton
UNIVERSITY of PENNSYLVANIA