

Shared control impairs cognitive control: Human responses inhibition slows when machines fail to inhibit

Patrick Bissett

Stanford University, Stanford, California, United States

Kriti Achyutuni

Stanford University, Stanford, California, United States

Jaime Ali H. Rios

Duke University, Durham, North Carolina, United States

Henry M. Jones

University of Chicago, Chicago, Illinois, United States

Russell Poldrack

Stanford University, Stanford, California, United States

Abstract

In order to fulfill goals, humans make use of cognitive control, which is a suite of processes to plan and manage thoughts and actions. One such process is response inhibition, which entails stopping a response when an action becomes inappropriate. Traditionally, response inhibition is measured in experimental settings in which humans have unilateral responsibility for inhibiting the action. However, in the real world, humans are increasingly sharing control with artificial intelligence (AI), with the paradigmatic case being partially automated vehicles. We designed an experiment that includes some aspects of partially automated vehicles and found that when humans share control with an AI that often but does not always stop, human response inhibition is significantly slowed even when the AI does not intervene. This reveals a cost of sharing control to human cognitive control, suggesting that the benefits of partial automation should be weighed against the costs of impaired human control.