Provably-safe interventions for Human-Cyber-Physical Systems (HCPS)

Sam Burden

Assistant Professor

Electrical Engineering

University of Washington

Seattle, WA USA

http://faculty.uw.edu/sburden

CNS #1565529

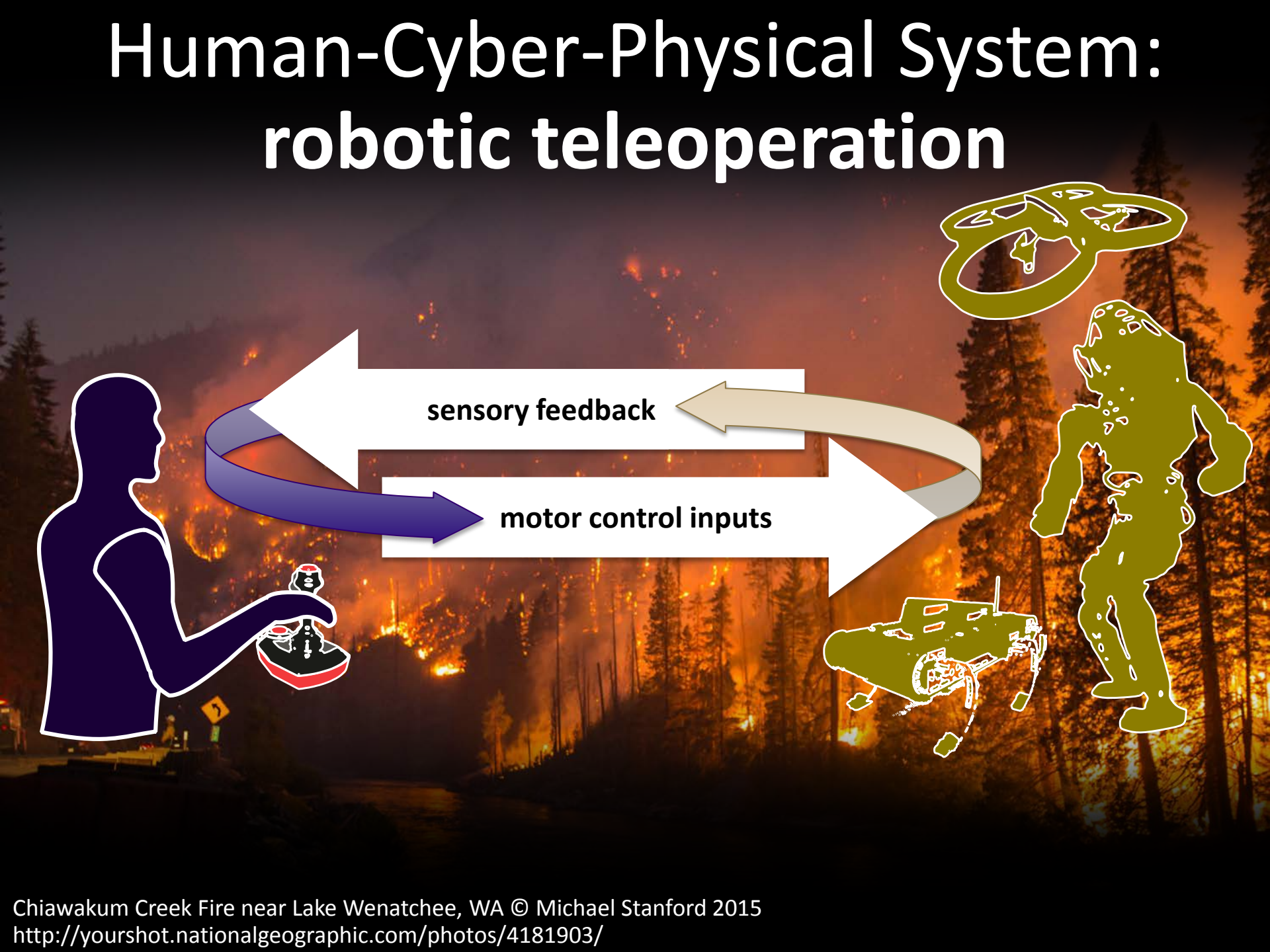
Eatai Roth

Darrin Howell

Momona Yamagami

Cydney Beckwith

Human-Cyber-Physical System: **robotic teleoperation**

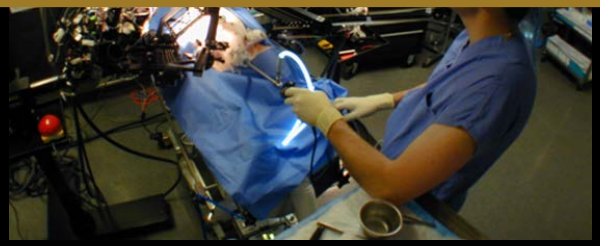
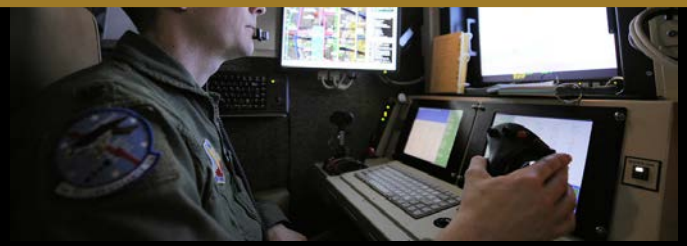


Chiawakum Creek Fire near Lake Wenatchee, WA © Michael Stanford 2015
http://yourshot.nationalgeographic.com/photos/4181903/

roles for humans and automation

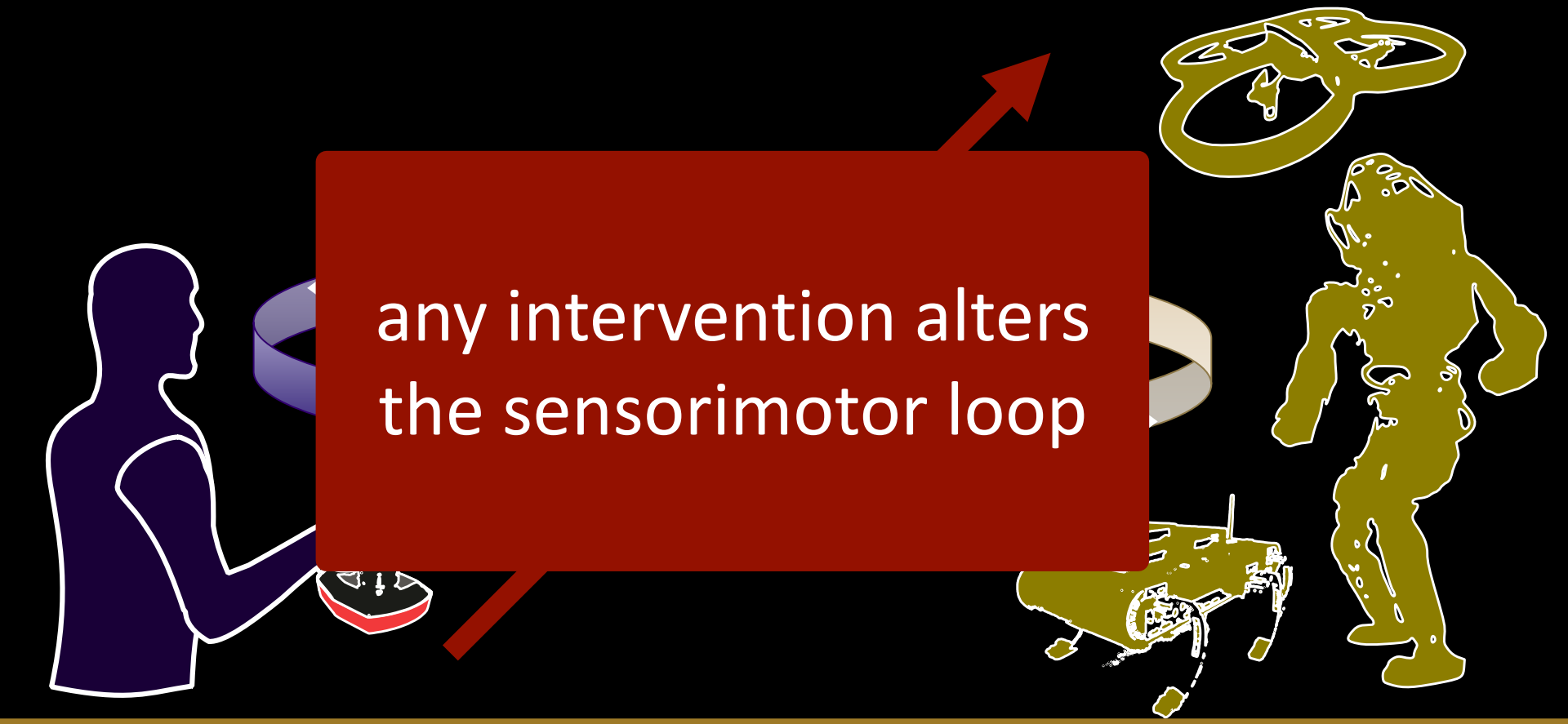


legal, **ethical**, and **political** concerns
ensure humans will remain in-the-loop


Nothwang, Robinson, Burden, McCourt, Curtis *IEEE Resilience Week* 2016
The Human Should be Part of the Control Loop?

intervening in Human-Cyber-Physical Systems



safe intervention requires validated predictive models for sensorimotor loops

predictable behavior from internal models

- theoretical and empirical evidence for pairing of **forward + inverse** models

  Bhushan, Shadmehr *Bio. Cybern.* 1999; Sanner, Kosha *Bio. Cybern.* 1999

forward model + **inverse** model

$$y = M(u) \qquad u = M^{-1}(y)$$

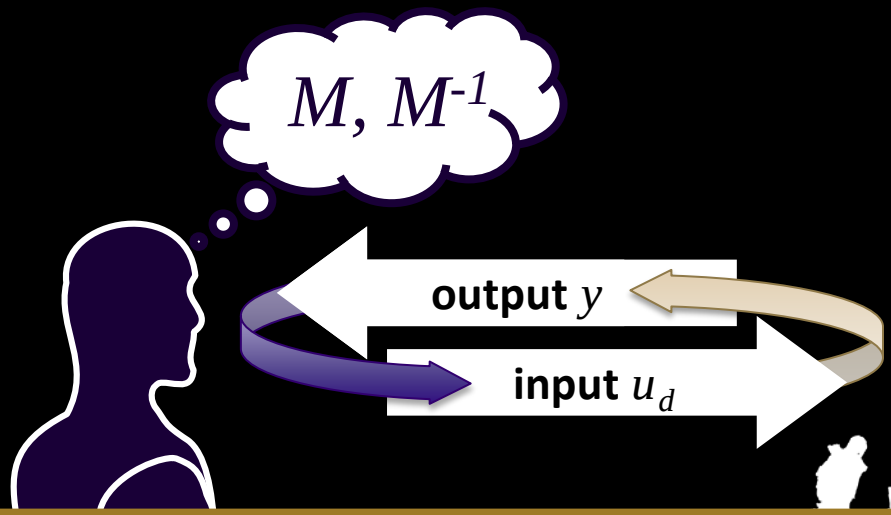
- parallels in control theory, robotics, artificial intelligence: adaptive control, internal model principle, learning

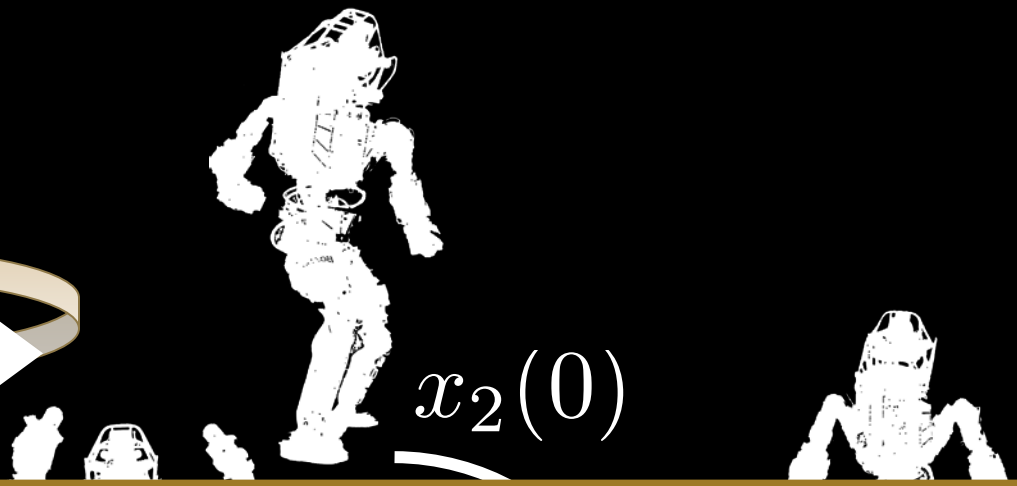
  Francis, Wonham *Automatica* 1976; Sastry, Bodson *Prentice Hall* 1989

  Sutton, Barto, Williams *IEEE CSM* 1992; Atkeson, Schaal *ICML* 1997

  Papavassiliou, Russell *IJCAI* 1999

theory for forward + inverse models



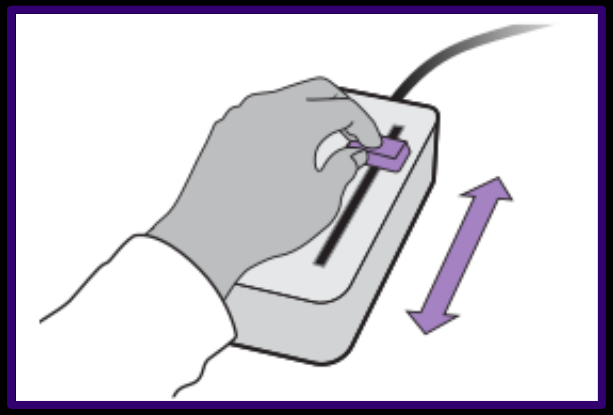


do humans learn forward + inverse models?

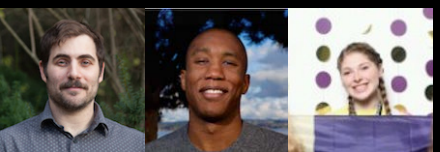
- **Theory results:**
  - for stable model pair, trajectories x_1 and x_2 converge to $\widehat{x}$
  - feedforward input "asymptotically inverts" dynamics

Robinson, Scobee, Burden, Sastry *SPIE-DSS* 2016
Dynamic inverse models in human-cyber-physical systems

experiments with forward + inverse models

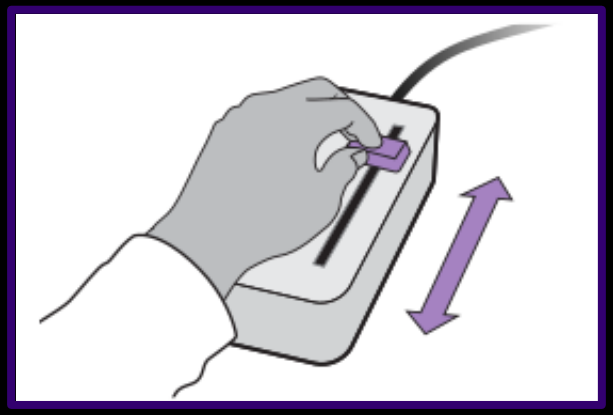


– subjects use 1-dimensional input device to control **cursor motion** to track **specified reference**

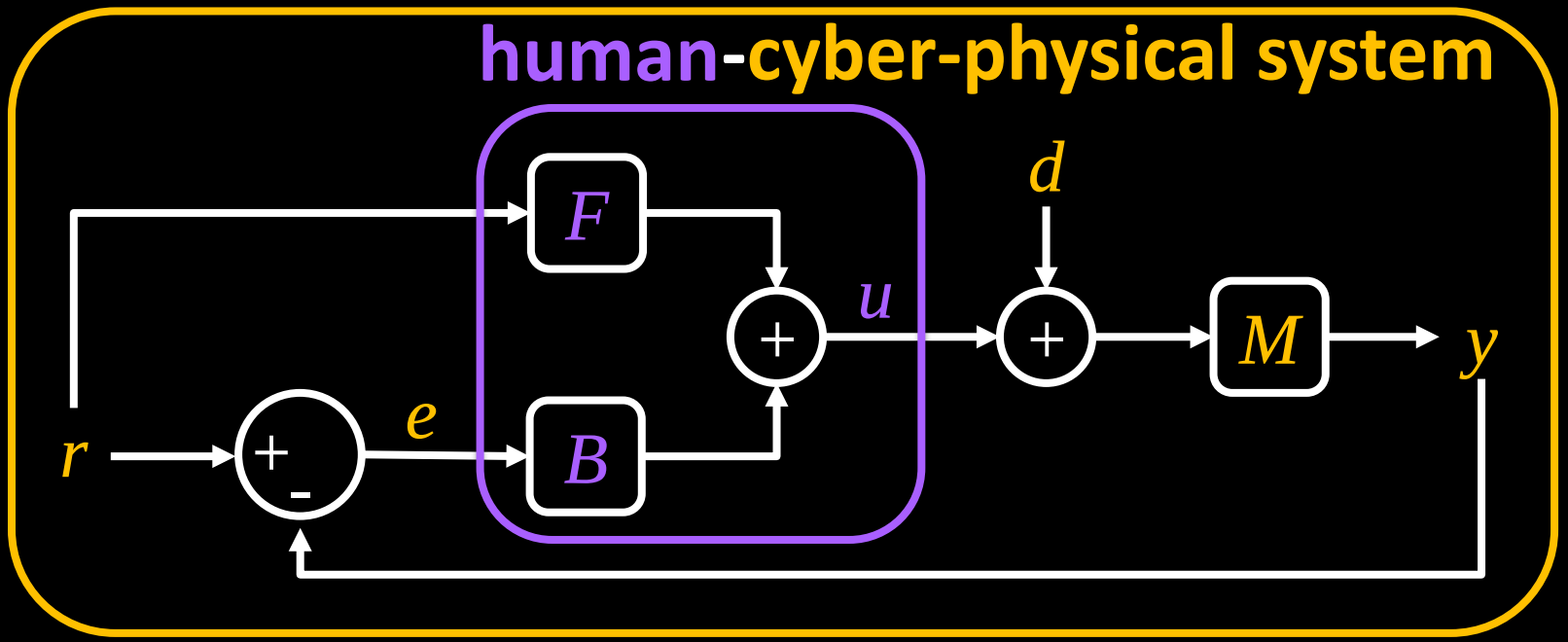


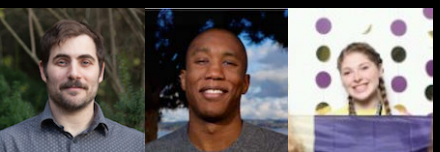
Roth, Howell, Beckwith, Burden *SPIE 2017*
Toward experimental validation of a model for human sensorimotor learning and control in teleoperation

experiments with forward + inverse models



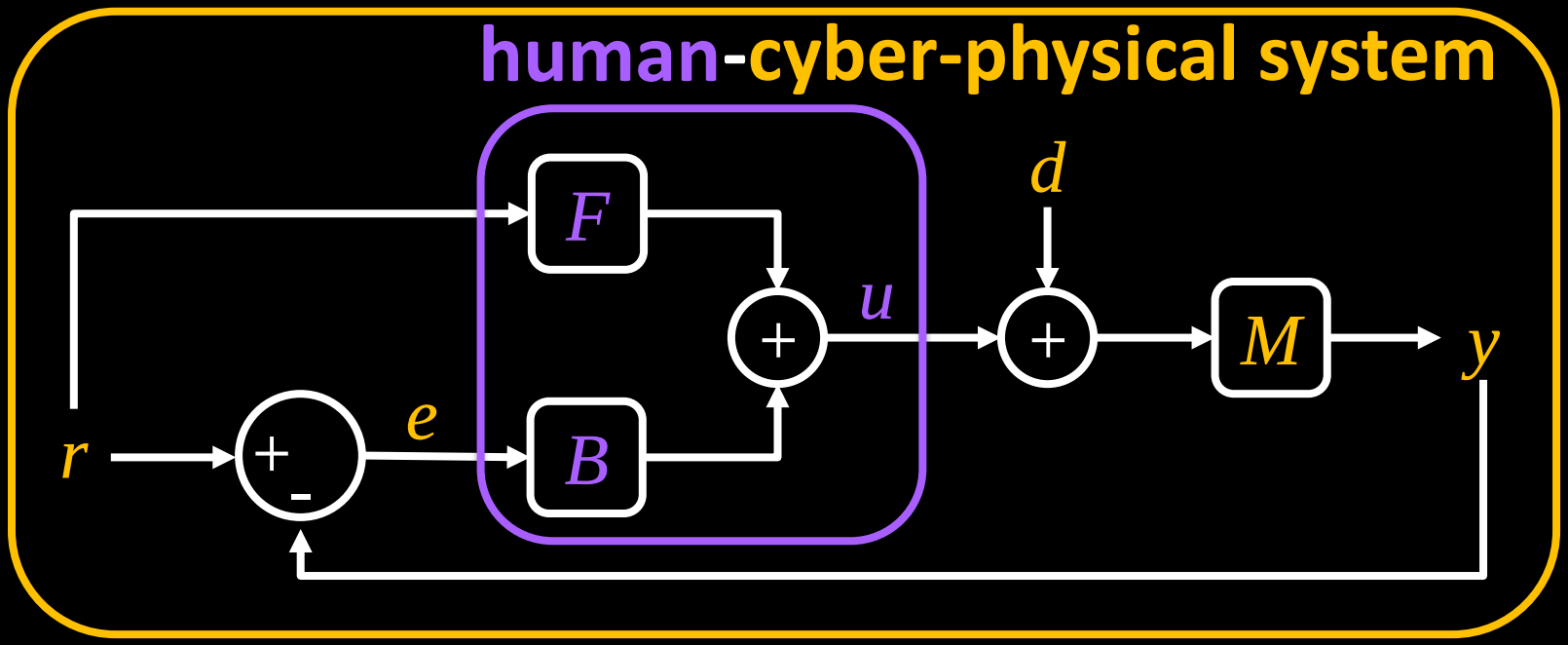
– subjects use 1-dimensional input device to control **cursor motion** to track **specified reference**



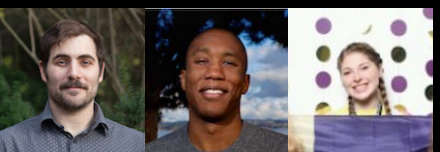


Roth, Howell, Beckwith, Burden *SPIE 2017*
Toward experimental validation of a model for human sensorimotor learning and control in teleoperation

empirically estimating learned model



- by varying reference (r) and disturbance (d), can estimate human feedforward (F), feedback (B)
- **human learns to invert specified model (M):** feedforward approximates the inverse ($F \approx M^{-1}$)



Roth, Howell, Beckwith, Burden *SPIE 2017*
Toward experimental validation of a model for human sensorimotor learning and control in teleoperation