mechanisms & consequences of attributing socialness to artificial agents

Emily S. Cross – University of Glasgow, Scotland

Artificial Intelligence & Society: Where are we headed? ERCEA, Brussels 26 October 2018
How do we make sense of others in a social world?
Challenge: Harness social robotics to advance our understanding of human social cognition

Social Cognition & Social Neuroscience

Robotics

Social Brain in Action Lab
SOCIAL ROBOTS in a nutshell

**Objective**: establish *behavioural & neural consequences* of social robot interaction

**Population**: European young adults

**Stream 1**: Ongoing

---

**Objective**: probe *malleability of socio-cognitive functioning* in young and advanced age

**Population**: European toddlers and older adults

**Stream 2**: Just now starting

---

**Objective**: explore *influence of cultural context* on perception & plasticity of “like me”-ness

**Population**: Japanese young adults

**Stream 3**: Watch this space!

---

**METHODS**

- mimicry
- functional neuroimaging
- neurostimulation
- training
... and Robots!
Snapshot of Ongoing Empirical Work

1. Automatic Imitation
2. Collaboration
3. Shared Representations
4. Empathy For Pain: Long-Term Social Intx & Modulation Of The Pain Matrix
5. Action Synchrony & Social Reward
Exploring Shared Representations with Robots

To what extent do we use similar neurocognitive mechanisms for social engagement with robots as we do with humans?

Empathy for Robots

Human Pain  Human Pleasure  Robot Pain  Robot Pleasure
Empathy for Robots

Each day of the robot socialising intervention comprised:

1. Free human-robot interaction
2. Games with the robot
3. Free play of the robot

Day 1 2 3 4 5
fMRI fMRI

Cross et al. (under revision)
Empathy for Robots

No evidence of neural mechanisms of empathy showing more overlap after socializing with robot

Cross et al. (under revision)
Empathy for Robots: Pursuing New Methods

**Standard analysis: Level of activation**

- Pre: Human > Robot
- Post: Human > Robot
- Neural overlap: Human > Robot

**Novel analysis: Shared representation**

- Training data: Robot pain > Robot pleasure
- Test data: Human pain > Human pleasure

**Brain structures and regions**

![Brain image](image-url)

**Social Brain in Action Lab**
Shared Representations
Shared Representations

Person Perception Network

Theory-of-Mind Network

ISVM

Happy

Angry

Sad

ISVM

Happy

Angry

Sad

Hortensius & Cross (in prep)
SOCIAL ROBOTS & AI: Into the Future
Upcoming Phil Trans Theme Issue (2019)

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

BIOLOGICAL SCIENCES

From Social Brains to Social Robots: Applying Neurocognitive Insights to Human-Robot Interaction

Editors: E. Cross // A. Wykowska // R. Hortensius
One of 3 primary subsections of AI, robotics is already reshaping manufacturing, military, and construction industries.

**Social robotics, per se,** is poised to dramatically change the industries and domains traditionally thought of as strictly human, including education, healthcare, services, communication, and sex.
SOCIAL ROBOTS & AI

How best to capitalize upon this potential, while avoiding pitfalls?

Interdisciplinarity to ensure crosstalk between experts across domains

Openness to updating our understanding of human-AI relationship

Clear delineation of boundaries between humans and machines (i.e., are the jobs/tasks for which we would never want to robots to enter?)
“We’re only barely scratching the surface of the brain’s social algorithms, which become even more complicated and unpredictable when we interface with technology.”

– Erik Sofge
Many Thanks!

Social Robots Team Members:
Bishakha Chaudhury
Lina Davitt
Anna Henschel
Ruud Hortensius
Te-Yi Hsieh
Laura Jastrzab
Henry Powell
Katie Riddoch
Rebecca Smith

Social Brain in Action Lab

www.so-bots.com
www.soba-lab.com
@brain_on_dance

ESRC
Economic & Social Research Council

ESRC Wales Doctoral Training Partnership
Partneriaeth Hyfforddiant Doethurol Cymru ESRC

Leverhulme Trust

Scottish Graduate School of Social Science
Sgoil Cheumnaichean Saidheans Sùisealta na h-Alba