



**Fabio Paglieri, ISTC-CNR Roma**  
**Simone Duca, University of Bristol**



# The logic of consciousness?

From CNCC to LogICCC, and back again

**E**UROPEAN  
S  
SCIENCE  
F  
FOUNDATION



# A tale of two talks

- The **CNCC** programme: a short introduction
- The **CONTACT** project: another short introduction
- *Scientific highlights* from CONTACT: possible links with LogICCC
- *Practical wisdom* from CONTACT: networking experiences
- The **METACOGNITION** project: yet another short introduction
- *Scientific highlights* from METACOGNITION: possible links with LogICCC
- *Practical wisdom* from METACOGNITION: networking experiences



# Basics of CNCC

- **CNCC** = *Consciousness in a Natural and Cultural Context*
  - *Large-scale research programme on consciousness, comprising 5 collaborative projects, which include 33 independent research groups across EU and beyond, from Nov 2006 to Nov 2009*
- **BASIC** - Brain, agency, self, intersubjectivity and consciousness
- **Boundaries of Mind** - Unconscious boundaries of mind: Research into the extended mind hypothesis
- **CEWR** - The conscious experience of what is reachable: Neural, behavioural, cultural and philosophical aspects
- **CONTACT** - Consciousness in interaction: The role of the natural and social environment in shaping consciousness
- **METACOGNITION** - Metacognition as a precursor to self-consciousness: Evolution, development and epistemology

# Who's who in CNCC

## **METACOGNITION**

Joëlle Proust, Institut Jean-Nicod, Paris, France  
Johannes Brandl, Universität Salzburg, Austria  
Hannes Leitgeb, University of Bristol, UK  
Josef Perner, Universität Salzburg, Austria  
Bernard Renault, Université Paris, France  
John David Smith, State University of New York at Buffalo, USA  
Josep Call, MPI for Evolutionary Anthropology, Leipzig, Germany

## **Boundaries of Mind**

Tjeerd Jellema, University of Hull, UK  
Albert Postma, Utrecht University, the Netherlands  
Johan Wagemans, University of Leuven, Belgium

## **CEWR**

Yann Coello, Université de Lille, France  
Joan Lopez-Moliner, Universitat de Barcelona, Spain  
Angela Sirigu, Université Claude Bernard, Lyon, France  
Jeroen Smeets, Vrije Universiteit Amsterdam, NL  
Bernard Pachoud, Centre de Recherche en Epistémologie Appliquée, Paris, France  
Alan Wing, University of Birmingham, UK

## **CONTACT**

Cristiano Castelfranchi, ISTC-CNR, Roma, Italy  
Andy Clark, University of Edinburg, UK  
Susan Hurley † / Finn Spicer, University of Bristol, UK  
Enrico Rambaldi, ISPF-CNR, Milano, Italy  
Ed S. Tan, Universiteit van Amsterdam, The Netherlands  
Thomas Metzinger, Johannes Gutenberg - Universität Mainz, Germany

## **BASIC**

Andreas Roepstorff, University of Aarhus, Denmark  
Christopher Frith, University College London, UK  
Shaun Gallagher, University of Central Florida, Orlando, USA  
Anthony Jack, Washington University, St. Louis, USA  
Tatjana Nazir, Hôpital Lyon Université, France  
Marcus Raichle, Washington University, St. Louis, USA  
Dan Zahavi, University of Copenhagen, Denmark  
Vittorio Gallese, Università degli Studi di Parma, Italy  
Patrick Haggard, University College London, UK  
Evan Thompson, University of Toronto, Canada  
Kai Vogeley, University of Cologne, Germany

# The CONTACT project

- **CONTACT** = *Consciousness in interaction: The role of the natural and social environment in shaping consciousness*
- Our research is built on the assumption that the brains and bodies of cognitive agents are **shaped by dynamical interactions** with both their *natural and social environments*
- Thus we *question* the assumption that conscious experience can be explained by the brain by itself, as opposed to the ***embodied brain in interaction with environments***, both natural and social



# The CONTACT project

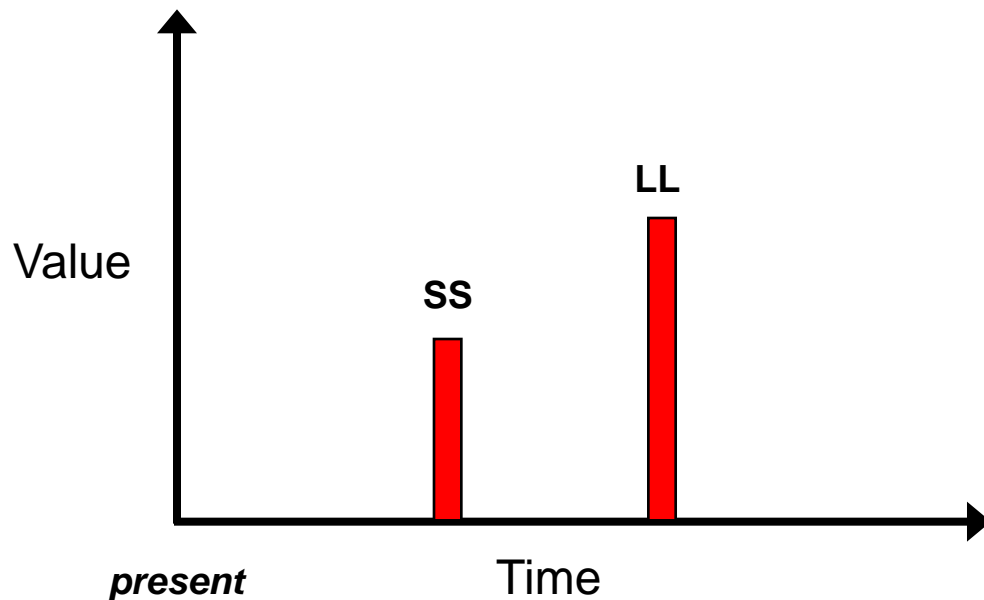
|                                                                                                            |                        |                                                           |
|------------------------------------------------------------------------------------------------------------|------------------------|-----------------------------------------------------------|
| IP-01: The Social Dimension of Consciousness                                                               | Castelfranchi, Roma    | <b><i>Social origins of conscious thought</i></b>         |
| IP-02: Making Up Minds: Sensorimotor Dynamics, Social Cognition, and Consciousness                         | Hurley/Spicer, Bristol | <b><i>Sensorimotor approaches to consciousness</i></b>    |
| IP-03: Active Consciousness, Embodiment, and Sense of Self                                                 | Clark, Edinburgh       | <b><i>Externalist and embodied models of the self</i></b> |
| IP-04: Emotional Feelings and Subject-Object Relationships                                                 | Tan/Frijda, Amsterdam  | <b><i>Emotions and consciousness</i></b>                  |
| IP-05: The Conceptual Roots of Consciousness in Interaction: Mapping Consciousness in the European Culture | Rambaldi, Milan        | <b><i>Historical approaches to consciousness</i></b>      |
| AP-01: Functional, Intentional, and Phenomenal Layers in the Human Self-Model                              | Metzinger, Mainz       | <b><i>Functional analysis of the sense of self</i></b>    |

# IP-01: Social dimension of consciousness

- *Individual* and *social* cognition as **co-evolving**, both in phylogenesis and in ontogenesis
- Consequence: possible to *apply models and tools used to study social interaction* (GT, SCT, argumentation theories) to **individual cognition**
- *The self as a multitude*
- Focus: **Intertemporal choice and multiple selves**

# Intertemporal choice

- Two results potentially available at ***different times*** : one is *smaller and sooner (SS)*, the other is *larger and later (LL)*
- They are **comparable** (can be weighted against each other) and **mutually exclusive** (you can achieve only one)



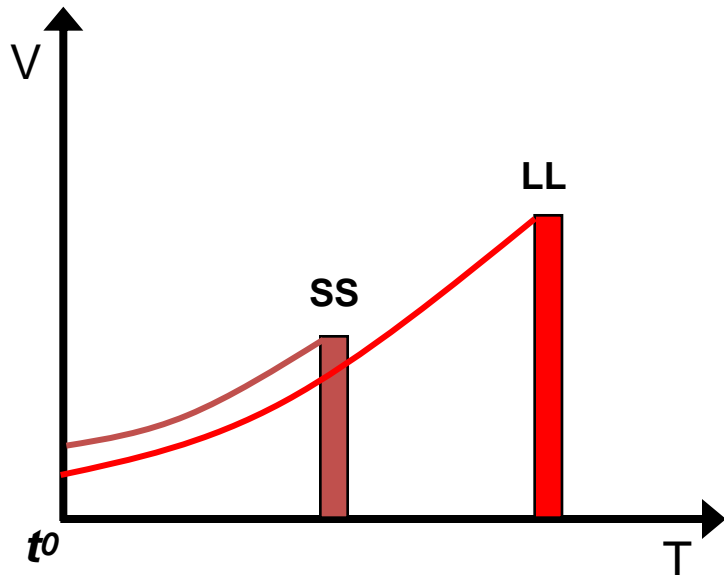
- Question: What should a rational agent prefer *now*?
- Fact: The answer depends on ***how the agent discounts utility over time***



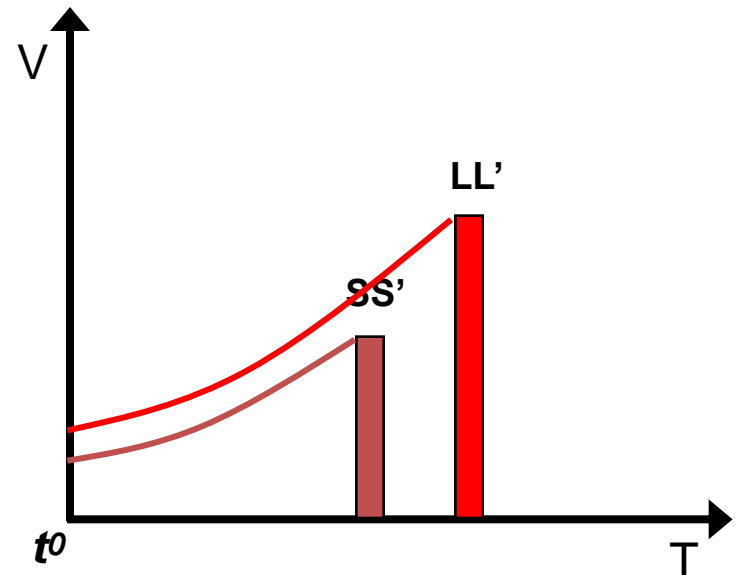
# Exponential discount and stable preferences

- Classic economics assumed **preference stability**: in the absence of external influences, the preferences of a rational agent should remain *consistent over time*
- This is guaranteed by **exponential time discounting**

**SS** > **LL** (always)

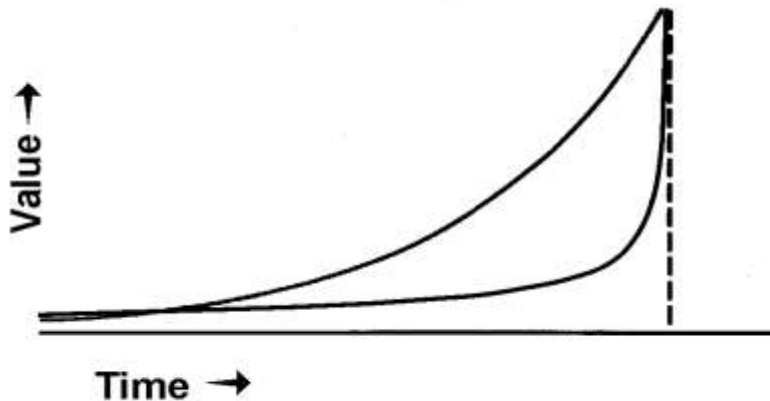


**LL'** > **SS'** (always)



# Hyperbolic time discounting

- Evidence in humans and other species suggests that future utility is discounted **hyperbolically**, not exponentially
- This implies that *motivational salience of rewards nearest in time is not proportional to that of later benefits*, so that expected utility is assessed **myopically**
- The curve is rather *flat when the delay is large*, whereas it becomes *very steep for short delays*



George Ainslie

# Hyperbolic discount and preference reversal

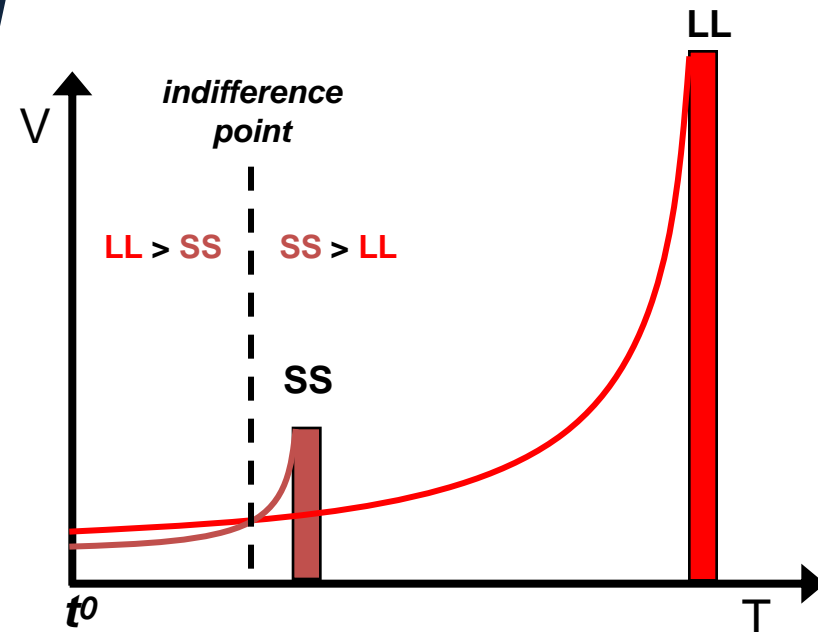
- Under certain conditions, hyperbolic discounting generates **time-dependent preference reversal**: *ceteris paribus*, the agent changes preferences *just because rewards closest in time appear disproportionately good*

**7\$ in 11 days > 5\$ in 10 days**

BUT

**5\$ today > 7\$ tomorrow**

*(similar with food and other goods)*



- Key variables: **impatience** factor, **distance** in time, **difference** in value at no delay

# Hyperbolic discounting and the will

- HD provides a natural explanation of *impulsive and short-sighted decisions* against one's own recognized best interest
- It also suggests that several forms of **akrasia**, or weakness of will (substance abuse, addiction, slacking, procrastination), exasperate standard HD: they are more acute forms of our temporal myopia, and this is why we *all suffer on occasion from these problems*
- So *what needs explaining*, if HD is true, is **willpower**, not its absence: How do we manage to overcome our short-sighted nature, thus achieving decent *intertemporal coordination* ?

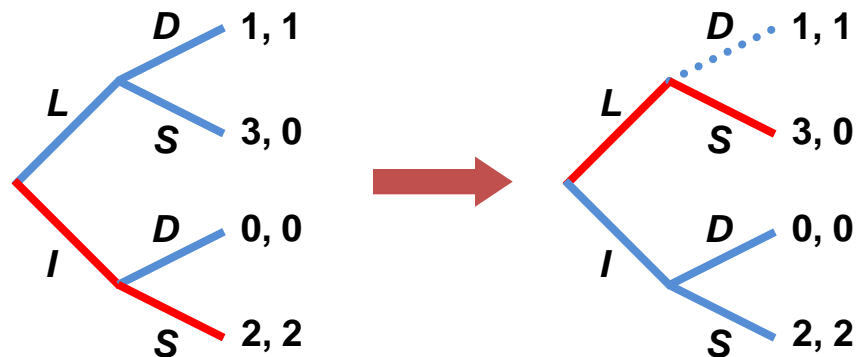
# Constraints theory: Elster and the sirens



- Present Self, *foreseeing* his future shift of motives and to avoid spoiling his long-term goal, **prevents his Future Self** from doing anything stupid
- Ulysses asks to be tied to the mast, to prevent himself from drowning when lured by the sirens



Liabile of game-theoretic interpretation in terms of *pruning the game tree in extended form*



# Intertemporal bargaining as IPD

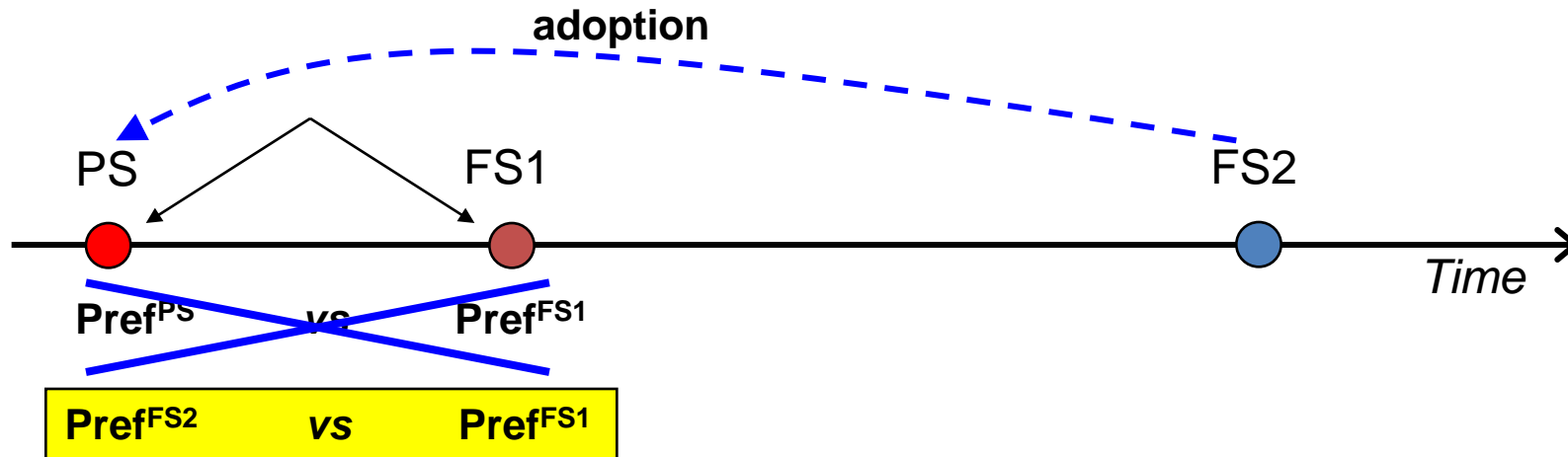
- «HD curves create a situation of *partial cooperation*, or limited warfare, among your successive motivational states. Their *individual interests in short-term rewards*, combined with their *common interests in stability of choice*, creates incentives very much like those in the much studied bargaining game, **repeated prisoner's dilemma**» (Ainslie 2001, p. 104)
- Ainslie is *vague* on the details, never gives the exact payoffs

|                              | later → LL<br><i>Cooperate</i> | later → SS<br><i>defect</i> |
|------------------------------|--------------------------------|-----------------------------|
| now → LL<br><i>cooperate</i> | 2, 2                           | 0, 3                        |
| now → SS<br><i>defect</i>    | 3, 0                           | 1, 1                        |

# Problems with game theory for ITB

- *Intrapersonal games*: problems of **transience**
  - *No retaliation/reward*, so no Tit-for-Tat (Bratman, 1999)
  - Need to *realize payoffs immediately* (Read, 2001)
- A possible solution with **intrapersonal preference adoption**: interacting *now* ( $t_0$ ) with my *future self* ( $t_1$ ) to maximize my payoffs at a *later* time ( $t_2$ )
- Bottom-line: intrapersonal bargaining is both **competitive and cooperative**

# Preference adoption



- **Preference adoption:** instead of playing with FS1 using his own preferences, *PS* uses the preferences of *FS2* and play with FS1 in order to *maximize FS2 utility*
- Instead of doing what I prefer now (at PS), **I do what I think I will prefer to have done afterwards** (at FS2): ‘thinking ahead’ and ‘making sacrifices for the future’
- This *circumvents* Read’s restriction, because by definition *FS2 payoffs are undetermined* until both PS and FS1 have moved



# CNCC Essay Award for Junior Scholar

- **2** prizes for essays on CNCC topics by junior scholars
- **1500 €** for each prize
- **44** eligible submissions
- **6** finalists
- **3** organizers
- **6** jurors
- **64** anonymous reviewers
- **89** independent reviews + **18** jurors reports
- (bad news) over **930** e-mails so far



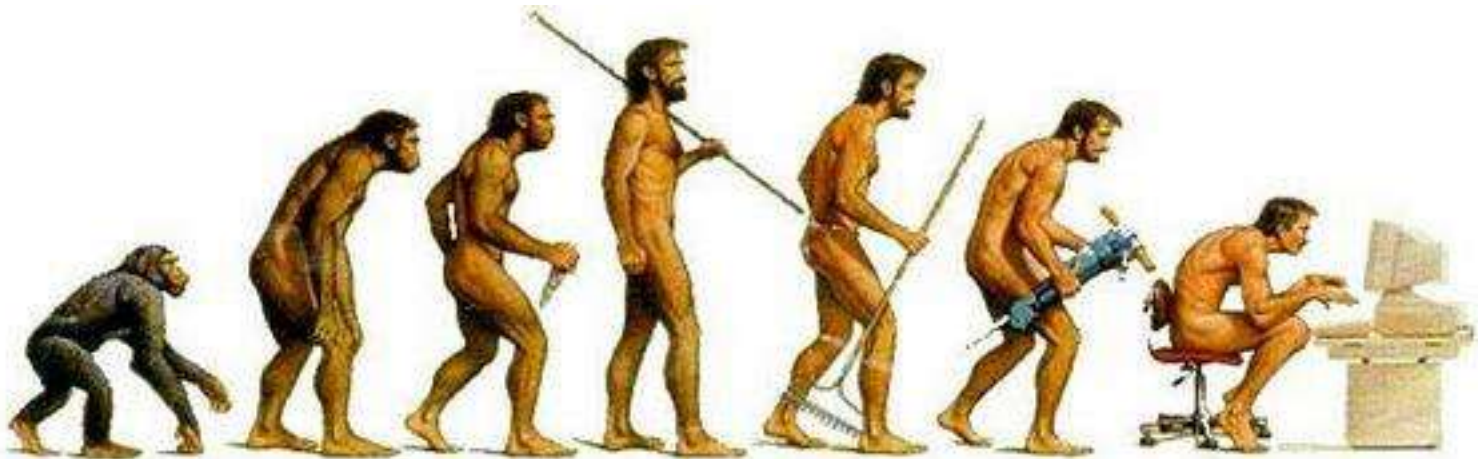
# Tough competition

- 6 finalists out of 44 submissions means a rate of acceptance close to **13%**, which is lower than most conferences for senior scholars
- For the 2 winners, this figure drops to **4.5%**, which is the acceptance rate of top scientific journals
- Very severe and strictly *blind peer-reviewing* by senior scholars
- Further independent assessment (also blind) by *jurors*



# The next challenge?

- To make this award scheme a **recurrent feature** of ESF-sponsored research programmes, using CNCC as a pilot
- Why might it be a **good idea**?
  - ESF-sponsored research programme are *interdisciplinary* and *thematically focused*, so allow for tough but coherent competition
  - Such an award scheme serves two purposes for the ESF: *dissemination* of the programme and *networking* among different research groups
  - Plus, with CNCC *it worked very well!*



# ESF Short Term Visits

- Both *intra-* and *inter-*programmes (LogICCC to/from CNCC)
- Priority given to *junior researchers*
- *Easy submission* (4-6 pp) and *quick processing* (2-3 months)
- Up to *6 weeks* stay
- *85 € / day* plus up to *500 €* for travel
- *Excellent cost/benefit ratio* for all parties involved (researchers, host institutions, ESF)



Personal experience: 1 month at the Phil Dept in Bristol working with *Hannes Leitgeb*, Feb 2008 – very rewarding & productive (now over to Simone, Hannes' PhD)