

MINDSHAPING AND MINDREADING IN SOCIAL COORDINATION

TAD ZAWIDZKI

**ASSOCIATE PROFESSOR & CHAIR,
PHILOSOPHY**

**CO-DIRECTOR, MIND/BRAIN INSTITUTE
GEORGE WASHINGTON UNIVERSITY**

HUMANS EXCEL AT SOCIAL COORDINATION

- **Cooperative projects like:**
 - Warfare
 - Alloparenting
 - Predator defense
 - Hunting
 - Regulation of sexual relations
 - Communication
 - (Re)creation of technologies
 - Social learning
- **No other mammal comes close**
- **Arguably the key to our relative biological success**
- **Often involves interaction with strangers (Seabright 2010)**

WHAT MAKES HUMAN SOCIAL COMPETENCE POSSIBLE?

- **Two stories:**
 - 1. Mindreading (the received view in psychology)**
 - We are far better natural psychologists than any other species
 - Our ability to read each other's minds, and, consequently, anticipate each other's behavior explains our virtuosity at social coordination
 - 2. Mindshaping (the alternative of which I'll try to convince you)**
 - Figuring out what's *really* on others' minds is intractable and irrelevant to successful social coordination
 - Instead, our minds are socially shaped from birth to facilitate complex social coordination in the absence of accurate mindreading

MINDREADING

- **In the classical philosophical understanding, this involves the attribution of concrete, unobservable states of a causal nexus (the “mind”), with causal relations to stimuli, other mental states, and behavior**
 - Sellars (1957); Lewis (1966); Armstrong (1968); Fodor (1987)
 - Sellars’s “myth of Jones”: birth of metaphor of folk as theoretical scientists
- **Huge influence on developmental and comparative psychology of social cognition**
 - Gopnik (1996); Gopnik and Wellman (1992)

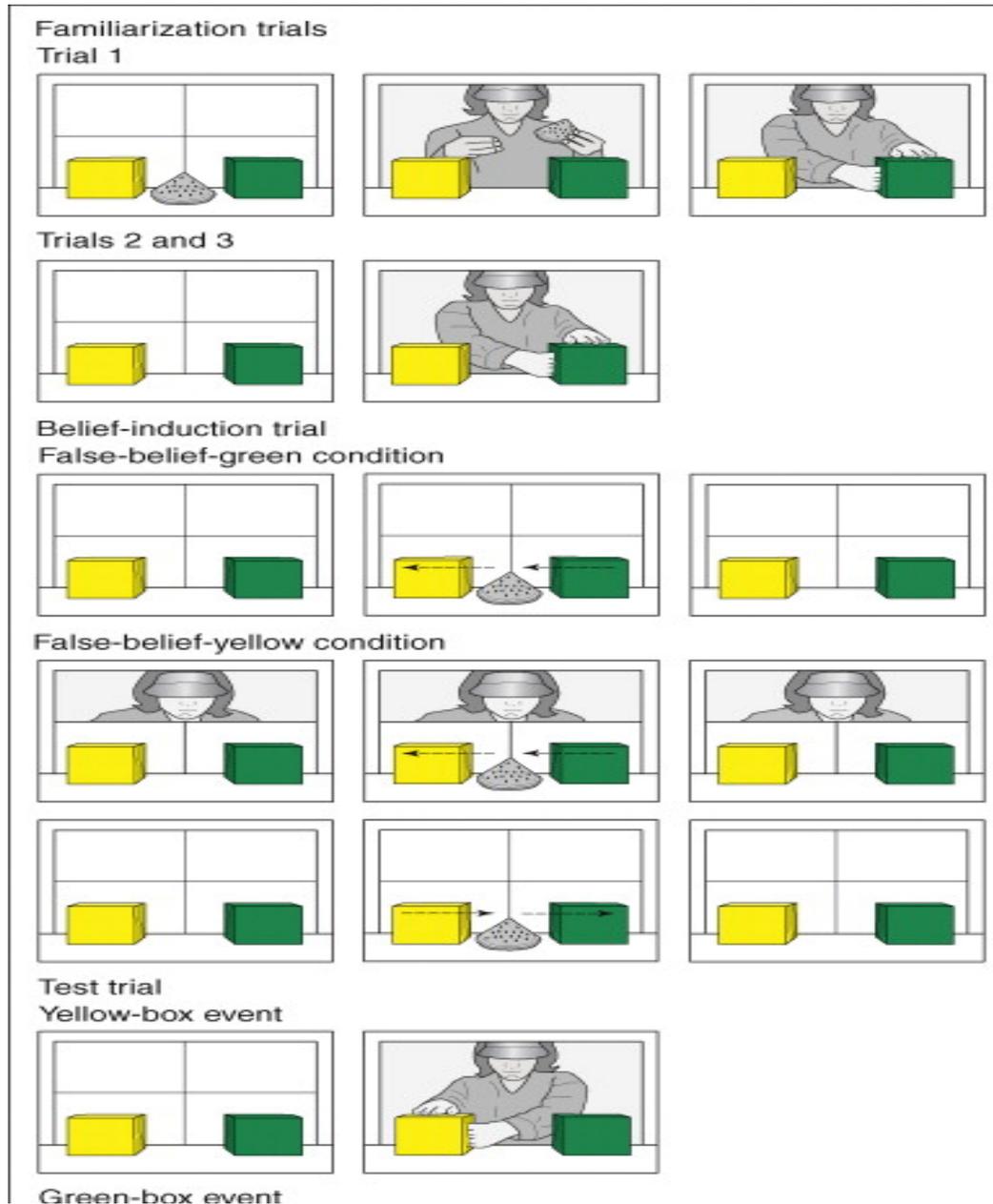
MINDREADING AND SOCIAL COORDINATION

- **Lewis (1969):**
 - We may achieve coordination by acting on our concordant expectations about each other's actions. And we may acquire those expectations ... by putting ourselves in the other fellow's shoes, to the best of our ability. If I know what you believe about matters of fact that determine the likely effects of your alternative actions, and if I know your preferences among possible outcomes and I know that you possess a modicum of practical rationality, then I can replicate your practical reasoning to figure out what you will probably do, so that I can act appropriately. ... In the case of a coordination problem ... one of the matters of fact that goes into determining the likely effects of your alternative actions is my own action. In order to figure out what you will do by replicating your practical reasoning, I need to figure out what you expect me to do. ... To replicate your reasoning, I may have to replicate your attempt to replicate my reasoning. ... This is not the end. I may reasonably expect you to realize that, unless I already know what you expect me to do, I may have to try to replicate your attempt to replicate my reasoning. So I may expect you to try to replicate my attempt to replicate your attempt to replicate my reasoning. So my own reasoning may have to include an attempt to replicate your attempt to replicate my attempt to replicate your attempt to replicate my reasoning. And so on. (27–28)
- **Kovács et al. (2010):**
 - Humans are guided by internal states such as goals and beliefs. Without an ability to infer others' mental states, society would be hardly imaginable. Social interactions, from collective hunting to playing soccer to criminal justice, critically depend on the ability to infer others' intentions and beliefs. Such abilities are also at the foundation of major evolutionary conundra. For example, the human aptitude at inferring mental states might be one of the crucial preconditions for the evolution of the cooperative social structure in human societies. (1830)
- **Often seen as key to human evolution**
 - Humphrey, 1980; Tooby & Cosmides, 1995, xvii; Baron-Cohen, 1999; Leslie, 2000, p. 61; Mithen, 2000; Sperber, 2000; Dunbar, 2000, 2003, 2009; Siegal, 2008, p. 22

THEORIES OF MINDREADING

- **Theory-theory (Gopnik & Wellman)**
 - Take theoretical scientist metaphor very seriously
 - Explicit theoretical inferences about unobservable causes of behavior lead children through trajectory of “scientific revolutions”
 - Culminate in understanding of false belief at age 4
- **Module-theory (Leslie, Baillergeon, Kovács)**
 - Chomsky-style innate competence gradually comes on-line as domain-general capacities like attention and working memory mature
 - Makes sense of infant competence at implicit false belief task
- **Simulation-theory (Goldman, Goodman, Heal, Harris)**
 - Using one’s own mind as model of others’ minds, while pretending to be “in their shoes”
- **Hybrids (Nichols & Stich 2003)**
 - Different aspects of mindreading handled by different capacities

FALSE BELIEF TASKS



WHAT ALL THEORIES OF MINDREADING SHARE

- **The point of it is the accurate attribution of mental states**
- **Mindreading can't help social coordination unless it involves identifying the actual beliefs and preferences that cause behavior**
- **The mental states we attribute when coordinating are taken to be real, concrete, unobservable causes of behavior, like viruses**
- **These are inferred from publicly observable behavior**

SOME KEY FEATURES OF BELIEFS AND PREFERENCES

- **Unobservable**
- **Concrete/causal**
 - “Real” causes of behavior, like viruses are real causes of symptoms
- **Behavioral appearance / mental reality distinction**
 - Just as with viruses: indistinguishable behaviors may have different mental causes
- **Bear propositional contents**
 - Opacity: how contents are described matters
- **Holism**
 - Beliefs & preferences triggered in response to stimuli, and cause observable behavior, only in indefinitely large blocks

PROBLEMS FOR RECEIVED VIEW

- **If human social coordination is supposed to depend on mindreading *in this sense* (accurate attribution of actual, concrete causes of behavior, that bear opaque, propositional contents, and interact holistically), then the received view is highly implausible**
 1. Humans are expert coordinators at least from age 1 (joint attention/intention), but show no appreciation for behavioral appearance / mental reality distinction, or opacity until much later (long after school starts)
 2. Many mindreading accounts of coordination rely on recursive attribution of mental states, but little evidence that children, and even most adults are good at attributing these efficiently and accurately
 3. Holism makes accurate mindreading *in this sense* computationally intractable
- **If this is required for human social coordination, its evolution is completely mysterious...**

MINIMAL MINDREADING?

- **The teleological stance (Gergely & Csibra 2003)**
 - Evidence that infants as young as 8 months interpret behavior as most rational/efficient option in a set of possible routes to a goal
- **The intentional stance (Dennett 1987)**
 - Like teleological stance, but adds informedness: same goal can lead to different behaviors if agents are differentially informed
 - No appearance/reality distinction: no further fact about which goals and informedness than what helps predict behavior on rationality assumption
- **Minimal theory of mind (Butterfill & Apperly 2013)**
 - A lot like intentional stance: agent behavior predicted based on goals and object “registrations”
 - Explains infant false belief tracking without assuming appreciation of opacity
 - But “goals” and “registrations” conceived of as “causal factors” intervening between stimuli and behavior.

TELEOLOGICAL REASONING IN INFANCY

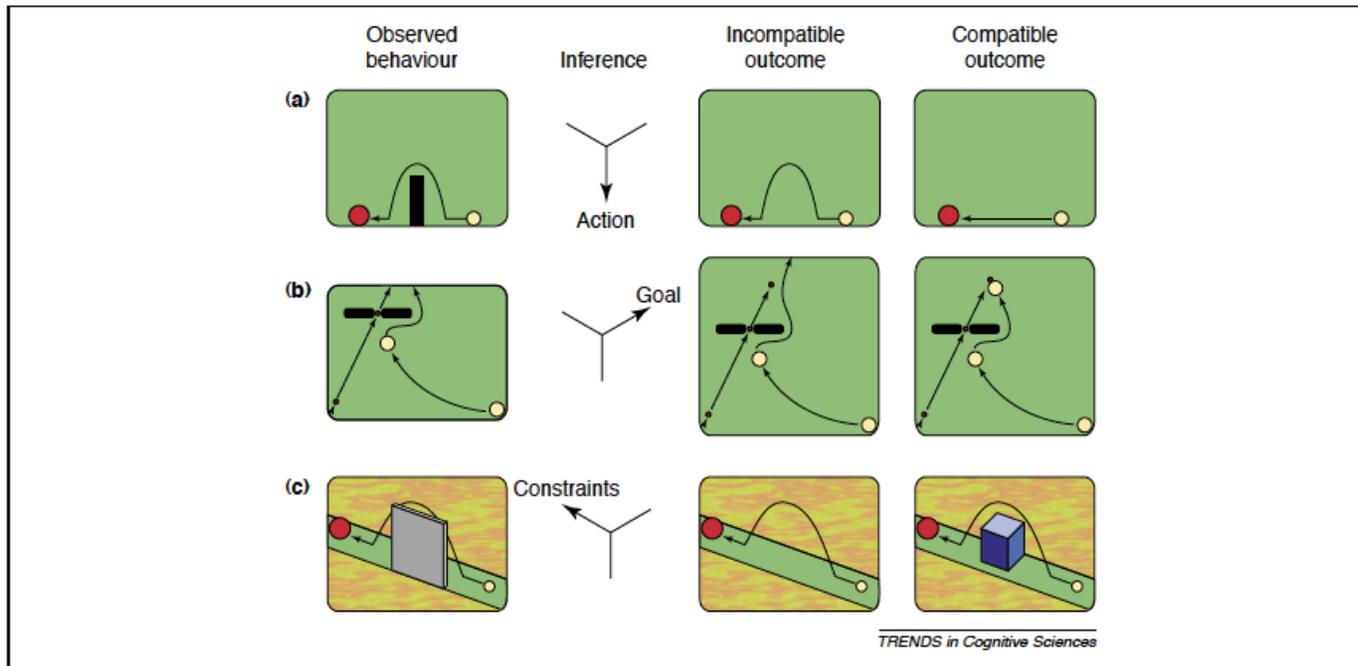


Fig. 1. Three types of inference that infants can draw based on a teleological representation of actions. One-year-old infants were habituated to the event depicted in the first column (Observed behaviour). Their interpretation of this event was tested by presenting them with two different outcomes, one of them being incompatible (second column), the other one being compatible (third column) with a possible inference based on a teleological representation of the event. Infants looked longer at the incompatible outcome than the compatible outcome events, indicating that they had made the assumed inference. (a) From the study in Ref. [7] and Experiment 1 in [10]. (b) Experiment 1A in [13]. (c) Experiment 2 in [13].

<http://tics.trends.com>

Gergely & Csibra (2003)

HOLISM

- **A step in the right direction, but still suffers from holism problem**
- **Indefinitely many goal/information attributions are compatible with finite bouts of observable behavior**
- **Indefinitely many behavioral predictions are compatible with particular goal/information attributions**
- **Doesn't yet solve problem of computational tractability**

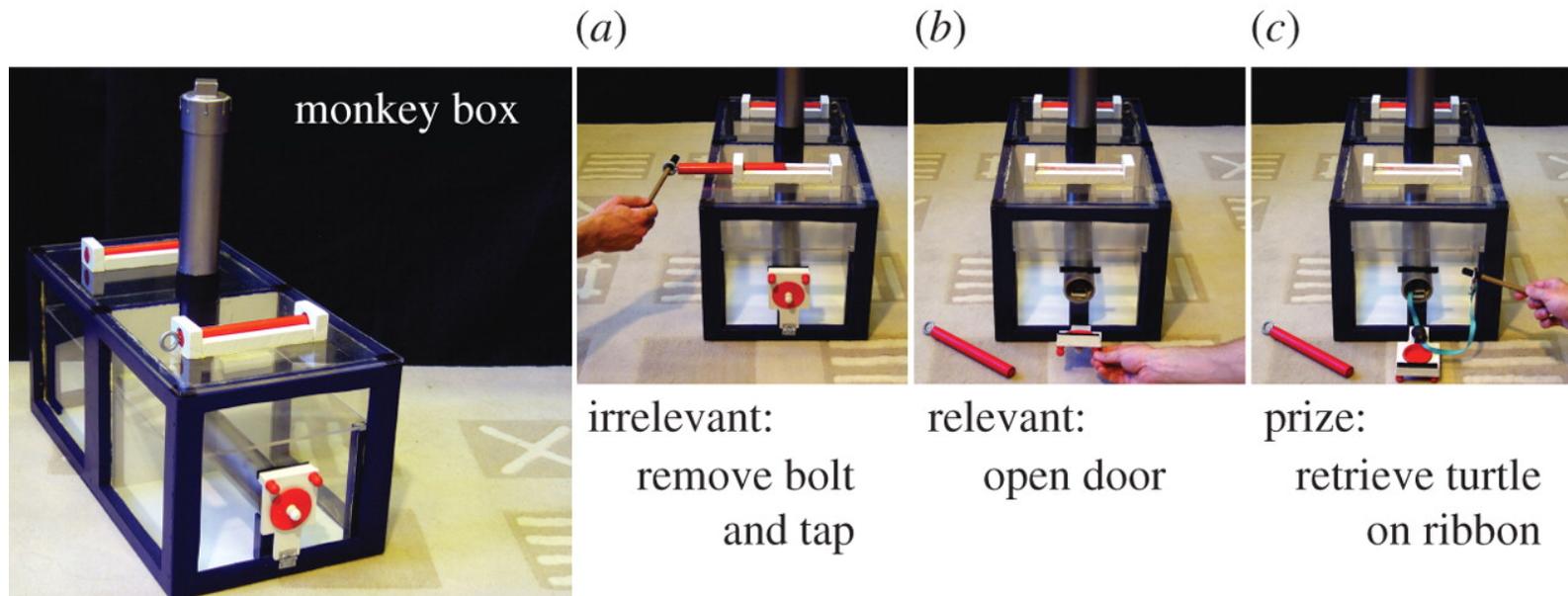
MINDSHAPING TO THE RESCUE?

- **Some proposed solutions to computational tractability problem are *externalist*:**
 - “Epistemic action” on cognitive domain can simplify it, making it more easily tractable by onboard cognitive resources (Clark 1997; 2003)
- **Epistemic action in the social realm: unreflective shaping of minds so that they are more alike**
- **Definition: A mindshaping mechanism is one the proper (i.e., evolved) function of which involves altering a behavioral disposition such that it approximates some (concrete or abstract) model**
 - Neutral on how it’s accomplished (mindreading may or may not contribute)
 - If these are sufficiently widespread, and models sufficiently homogenous in a population, then members should come to resemble each other more closely in behavioral responses to similar stimuli
 - Holism might be mitigatable (if not entirely avoidable):
 - One need consider only intuitively obvious goals, information, and reasoning strategies when interpreting others if one *happens to be* (*whether one knows it or not*) a product of similar mindshaping

VARIETIES OF MINDSHAPING

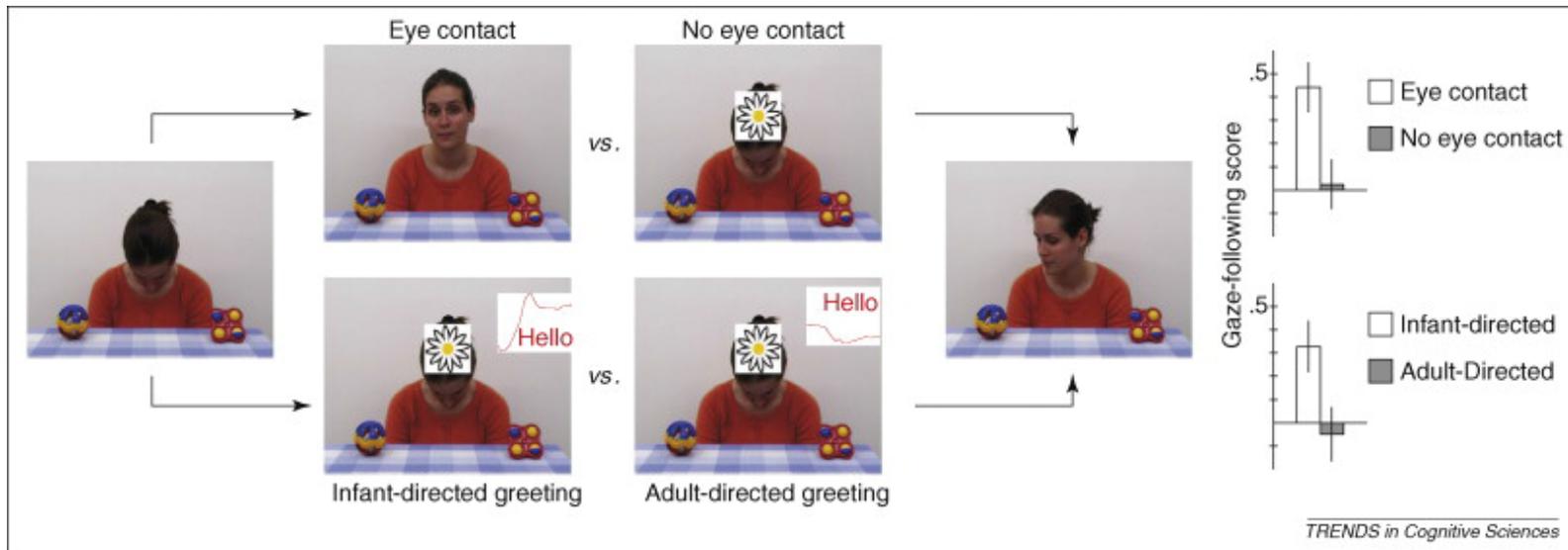
- Imitation of goals is very widespread in animal kingdom (Zentall 2006)
- Are there distinctively human kinds that are independent of sophisticated mindreading, and can explain human virtuosity at social coordination?
- Yes:
 1. Overimitation
 2. Natural pedagogy
 3. “Costly” norm institution/policing
 4. Narrative self-constitution

OVERIMITATION



Lyons, et al. 2011; see also: Legare & Nielsen 2015; Nielsen & Tomaselli 2010.

NATURAL PEDAGOGY



Csibra & Gergely 2009.

NORM INSTITUTION & ENFORCEMENT

- **Humans *seem* willing to pay high costs to enforce interaction norms**
 - E.g., ultimatum game data show enforcement of resource distribution norms at the cost of rejecting better-than-zero offers
- **Not necessarily costly given typical human preferences for socially sanctioned behavior**
 - Conformism bias (Henrich 2004)
 - Brain areas implementing negative reinforcement in conditioning active when judgment departs from population mean (Klucharev et al. 2009)
 - Social rejection shares somatosensory representations with physical pain (Kross et al. 2011)
- **But such preferences seem unique to humans and may help explain social coordination:**
 - Make mindshaping to respect coordination-promoting norms easier

NARRATIVE SELF- CONSTITUTION

- **Public language makes available ways of interpreting self and others that promote coordination**
- **Publicly encoded stereotypes, scripts, frames used to understand respective roles in typical interactions**
- **Loaded (self-) interpretive categories like “parent”, “teacher”, “soldier”, “judge”, “representative”, “owner”, “server”, “customer”, “European”, “American”, “Italian”, etc. are like software implemented on our brains, or interaction tools**
- **Explains why humans are obsessed with acquiring such interpretive categories**
- **Also explains why inaccurate ones can be successful when prevalent:**
 - Their raison d'être is promoting as much successful coordination as possible, so *their prevalence is all that matters*
 - Frequency-dependent phenomena, like language
 - Accuracy is beside the point, as long as human brains can implement them
 - Bibles sell well because Bibles sell well...

CAVEAT: MINDSHAPING PRESUPPOSES SOME KIND OF MINIMAL MINDREADING

- **Overimitation (Gergely et al. 2002 on rational imitation):**



- **Natural pedagogy: Gergely & Csibra speak of nested, proto-Gricean intentions...**
- **Narrative self-constitution: Doesn't linguistic communication presuppose mindreading?**
- **Obviously some kind of bootstrapping at work...**
 - Start with roughly the minimal mindreading we share with other primates
 - Add socio-ecology that incentivized widescale coordination on cooperative projects (Sterelny 2012 argues this characterized human prehistory)
 - Leads to initial mindshaping, which then makes more sophisticated mindreading reliable and possible, etc.

WHAT DOES THIS HAVE TO DO WITH GAME THEORY?

- And, in particular, Wynn Sterling's theory of conditional games?
- I'm EXTREMELY new to this, but here are my 2 cents:
- 2 construals:
 1. As model of actual reasoning process in real coordinators:
 - What of threat of holism and computational tractability in determining preferences of leaders, and others' takes on these preferences?
 - Mindshaping might help here: what matters is not actual preference, but preference leader is "supposed" to have given stereotypes according to which everyone's been shaped (whether they know it or not)
 - In humans, abstract models serve not just to describe, but also to regulate, and thereby, self-fulfill (McGeer 1996; 2007)
 2. As theorist-modeler's tool:
 - But then what is its goal?
 1. Modeling actual reasoning processes? See 1. above.
 2. Encoding norms of reasoning? Ross: not a normative project.
 3. Competence-level specification... Best bet.
 - Compare some insect that approximates optimal computational solution to travelling salesman problem...

FEARS FOR THE FUTURE

- **We can think of cultures as ecologies of interaction tools their members are mindshaped to implement**
 - Human social coordination depends on a certain degree of homogeneity (Skyrms 2003), and culture can be seen as an evolved response to this problem
 - Evidence that subjects who tap in rhythm more likely to cooperate on subsequent public goods game (Wiltermuth & Heath 2009)
- **But as Durkheim (1897) notes, modernity ushers in historically unprecedented heterogeneity of culture**
 - And he had no idea of the echo chambers the internet has created
- **It's possible that contemporary balkanization of coordinative ecologies, via the proliferation of increasingly heterogeneous interaction tools, i.e., (self-)interpretive frameworks like religions, ideologies, etc., may seriously threaten human coordinative capacity...**