



# Minimal Theory of Mind Joint Action

[s.butterfill@warwick.ac.uk](mailto:s.butterfill@warwick.ac.uk)

A black and white photograph of two young children, a boy and a girl, standing close together against a dark background. The boy, on the left, has light-colored hair and is wearing a patterned long-sleeved shirt under dark overalls. The girl, on the right, also has light-colored hair and is wearing a dark, patterned dress. They are both smiling and looking towards the camera.

# challenge

Explain the emergence, in evolution or development, of full-blown theory of mind cognition.



# challenge

Explain the emergence, in evolution or development, of full-blown theory of mind cognition.

# puzzle

Theory of mind *abilities* are widespread

**18-month-olds** point to inform, and predict actions based on false beliefs

(Liszkowski et al 2006)

(Onishi & Baillargeon 2005;  
Southgate et al 2007)

Theory of mind *abilities* are widespread

**18-month-olds** point to inform, and predict actions based on false beliefs

**Scrub-jays** selectively re-cache their food in ways that deprive competitors of knowledge of its location

(Liszkowski et al 2006)

(Onishi & Baillargeon 2005;  
Southgate et al 2007)

(Clayton, Dally & Emery 2007)

Theory of mind *abilities* are widespread

**18-month-olds** point to inform, and predict actions based on false beliefs

(Liszkowski et al 2006)

(Onishi & Baillargeon 2005;  
Southgate et al 2007)

**Scrub-jays** selectively re-cache their food in ways that deprive competitors of knowledge of its location

(Clayton, Dally & Emery 2007)

**Chimpanzees** conceal their approach from a competitor's view, and act in ways that are optimal given what another has seen

(Hare, Call & Tomasello 2006)

(Hare, Call & Tomasello 2001)

Theory of mind *abilities* are widespread

**18-month-olds** point to inform, and predict actions based on false beliefs

**Scrub-jays** selectively re-cache their food in ways that deprive competitors of knowledge of its location

**Chimpanzees** conceal their approach from a competitor's view, and act in ways that are optimal given what another has seen

(Liszkowski et al 2006)

(Onishi & Baillargeon 2005;  
Southgate et al 2007)

(Clayton, Dally & Emery 2007)

(Hare, Call & Tomasello 2006)

(Hare, Call & Tomasello 2001)

Theory of mind *abilities* are widespread

**18-month-olds** point to inform, and predict actions based on false beliefs

(Liszkowski et al 2006)

(Onishi & Baillargeon 2005;

# question

What could infants, chimps and scrub-jays represent that would enable them, within limits, to track others' perceptions, knowledge, beliefs and other propositional attitudes?

approach from a competitor's view, and act in ways that are optimal given what another has seen

(Hare, Call & Tomasello 2006)

(Hare, Call & Tomasello 2001)

Theory of mind *abilities* are widespread

**18-month-olds** point to inform, and predict actions based on false beliefs

**Scrub-jays** selectively re-cache their food in ways that deprive competitors of knowledge of its location

**Chimpanzees** conceal their approach from a competitor's view, and act in ways that are optimal given what another has seen

Representing perceptions, knowledge states and beliefs is hard

Theory of mind *abilities* are widespread

**18-month-olds** point to inform, and predict actions based on false beliefs

**Scrub-jays** selectively re-cache their food in ways that deprive competitors of knowledge of its location

**Chimpanzees** conceal their approach from a competitor's view, and act in ways that are optimal given what another has seen

Representing perceptions, knowledge states and beliefs is hard, for it requires

(a) conceptual sophistication

- takes years to develop
- development tied to acquisition of executive function and language
- development facilitated by training and siblings

Theory of mind *abilities* are widespread

**18-month-olds** point to inform, and predict actions based on false beliefs

**Scrub-jays** selectively re-cache their food in ways that deprive competitors of knowledge of its location

**Chimpanzees** conceal their approach from a competitor's view, and act in ways that are optimal given what another has seen

Representing perceptions, knowledge states and beliefs is hard, for it requires

(a) conceptual sophistication

- takes years to develop
- development tied to acquisition of executive function and language
- development facilitated by training and siblings

(b) scarce cognitive resources

- attention
- working memory

Theory of mind *abilities* are widespread

**18-month-olds** point to inform, and predict actions based on false beliefs

**Scrub-jays** selectively re-cache their food in ways that deprive competitors of knowledge of its location

**Chimpanzees** conceal their approach from a competitor's view, and act in ways that are optimal given what another has seen

Representing perceptions, knowledge states and beliefs is hard, for it requires

(a) conceptual sophistication

- takes years to develop
- development tied to acquisition of executive function and language
- development facilitated by training and siblings

(b) scarce cognitive resources

- attention
- working memory



# challenge

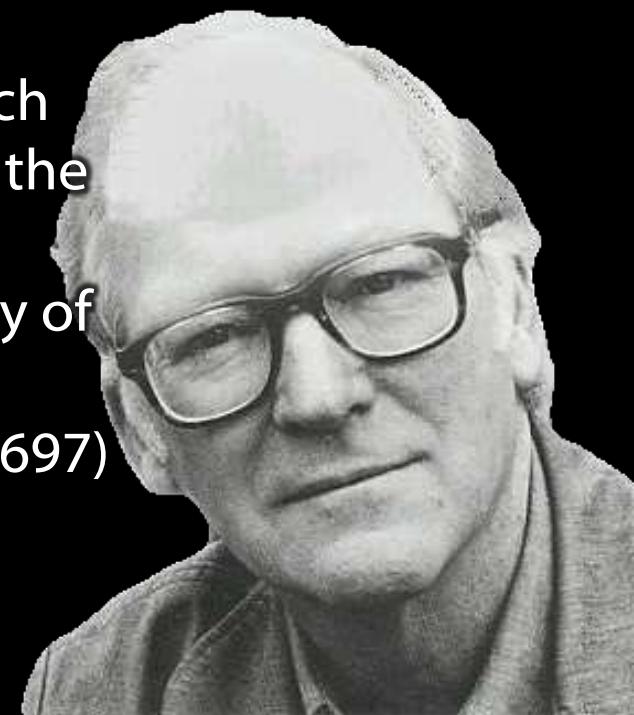
Explain the emergence, in evolution or development, of full-blown theory of mind cognition.

# puzzle

What could infants, chimps and scrub-jays represent that would enable them, within limits, to track others' perceptions, knowledge, beliefs and other propositional attitudes?

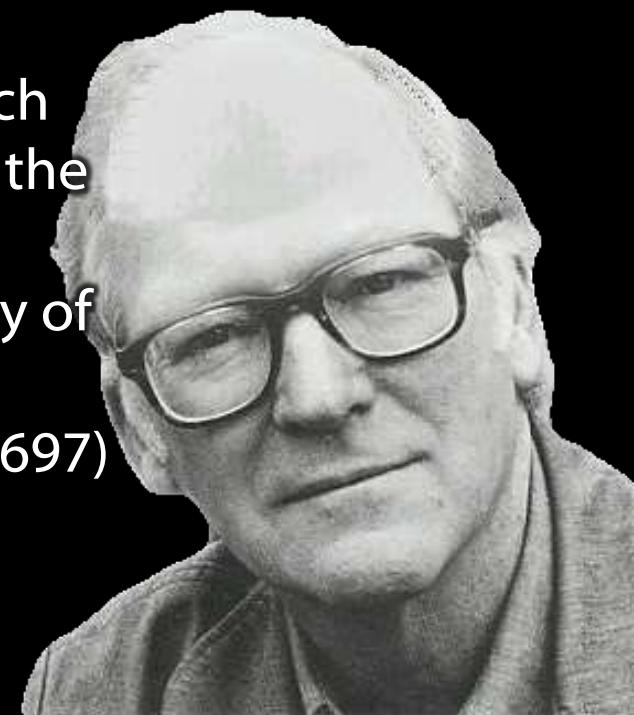
“We are stuck with our two main ways of describing and explaining things, one which treats objects and events as mindless, and the other which treats objects and events as having propositional attitudes. I see no way of bridging the gap”

(Davidson 2003:697)



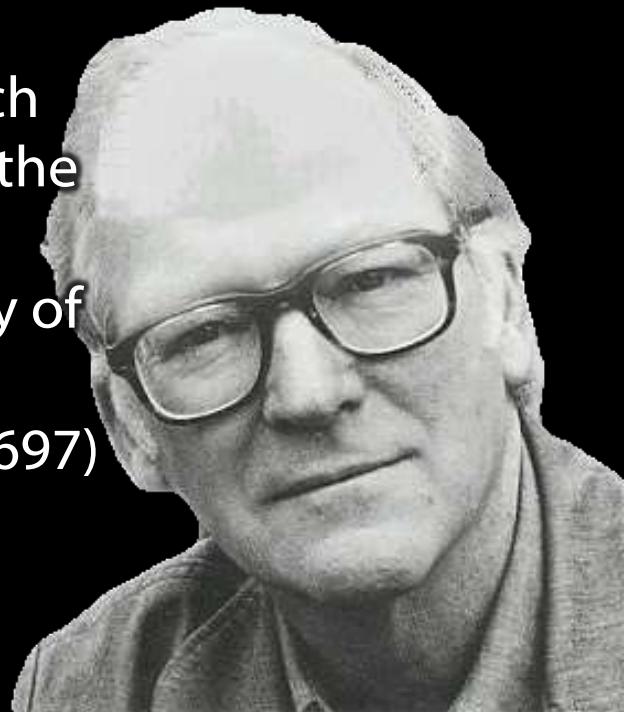
“We are stuck with our two main ways of describing and explaining things, one which treats objects and events as mindless, and the other which treats objects and events as having propositional attitudes. I see no way of bridging the gap”

(Davidson 2003:697)



“We are stuck with **our** two main ways of describing and explaining things, one which treats objects and events as **mindless**, and the other which treats objects and events as having **propositional attitudes**. I see no way of bridging the gap”

(Davidson 2003:697)





“We are stuck with our  
describing and explaining  
treats objects and events as  
other which treats objects as  
having propositional attitudes  
bridging the gap”

(Davidson)

Your *field* = a set of  
objects related to you by  
proximity, orientation,  
lighting and other factors



Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors



*proximity*

*orientation*

*lighting*

*barriers*

*trajectory*

Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

You *encounter* an object = it is in your field



*proximity*

*orientation*

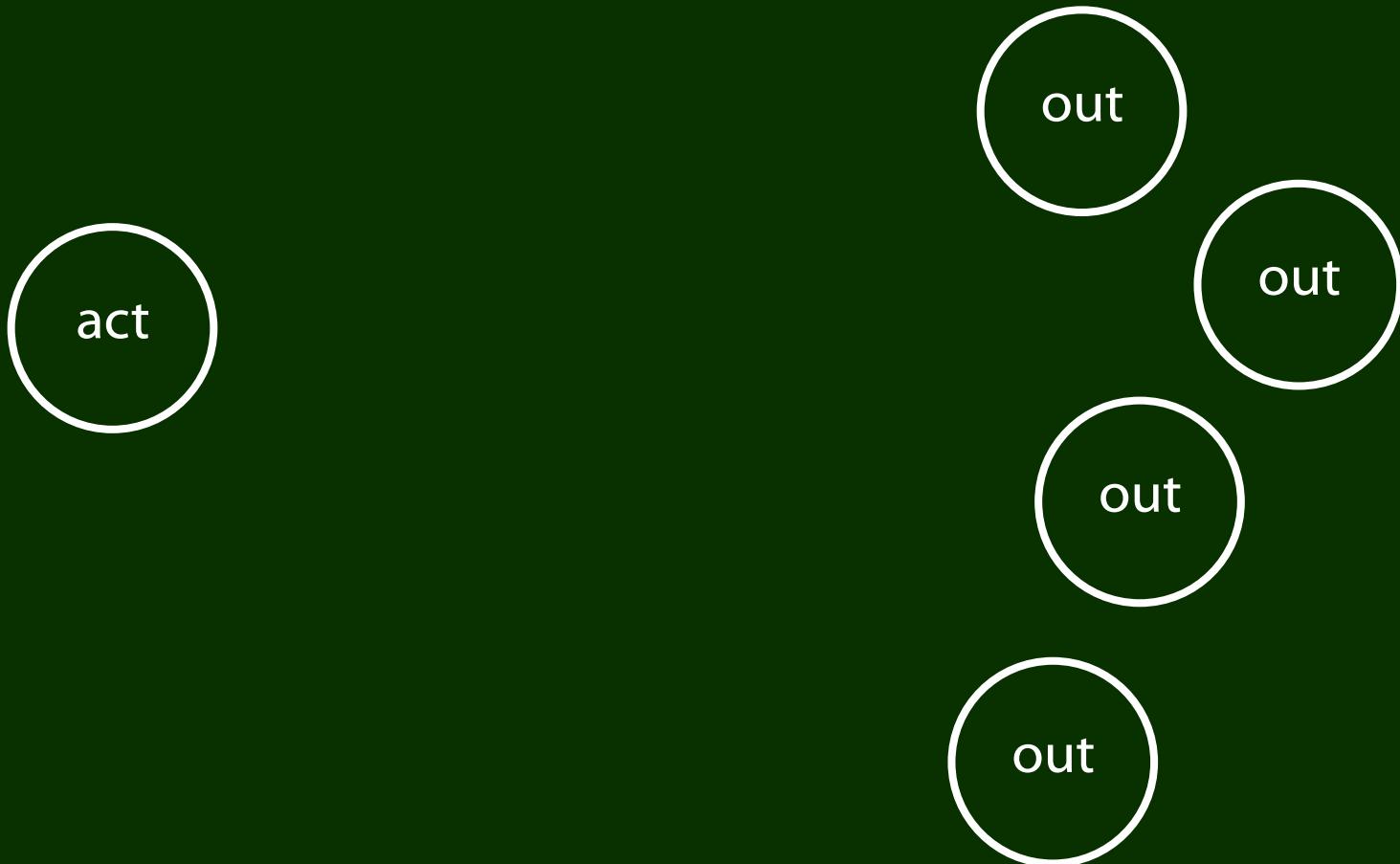
*lighting*

*barriers*

*trajectory*

Detour  
Goals are not intentions

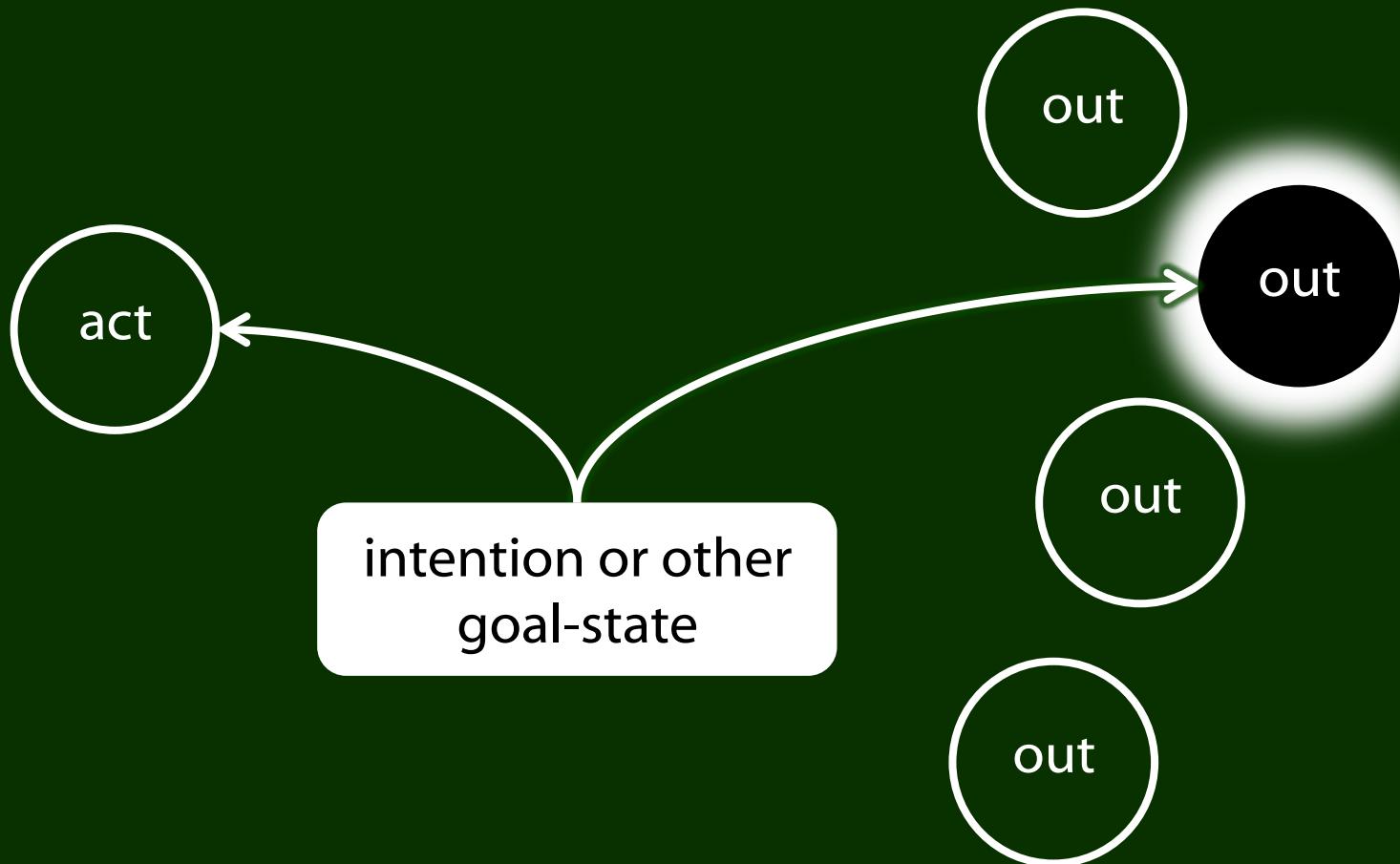
What is the relation between an action and the goal (or goals) to which it is directed?



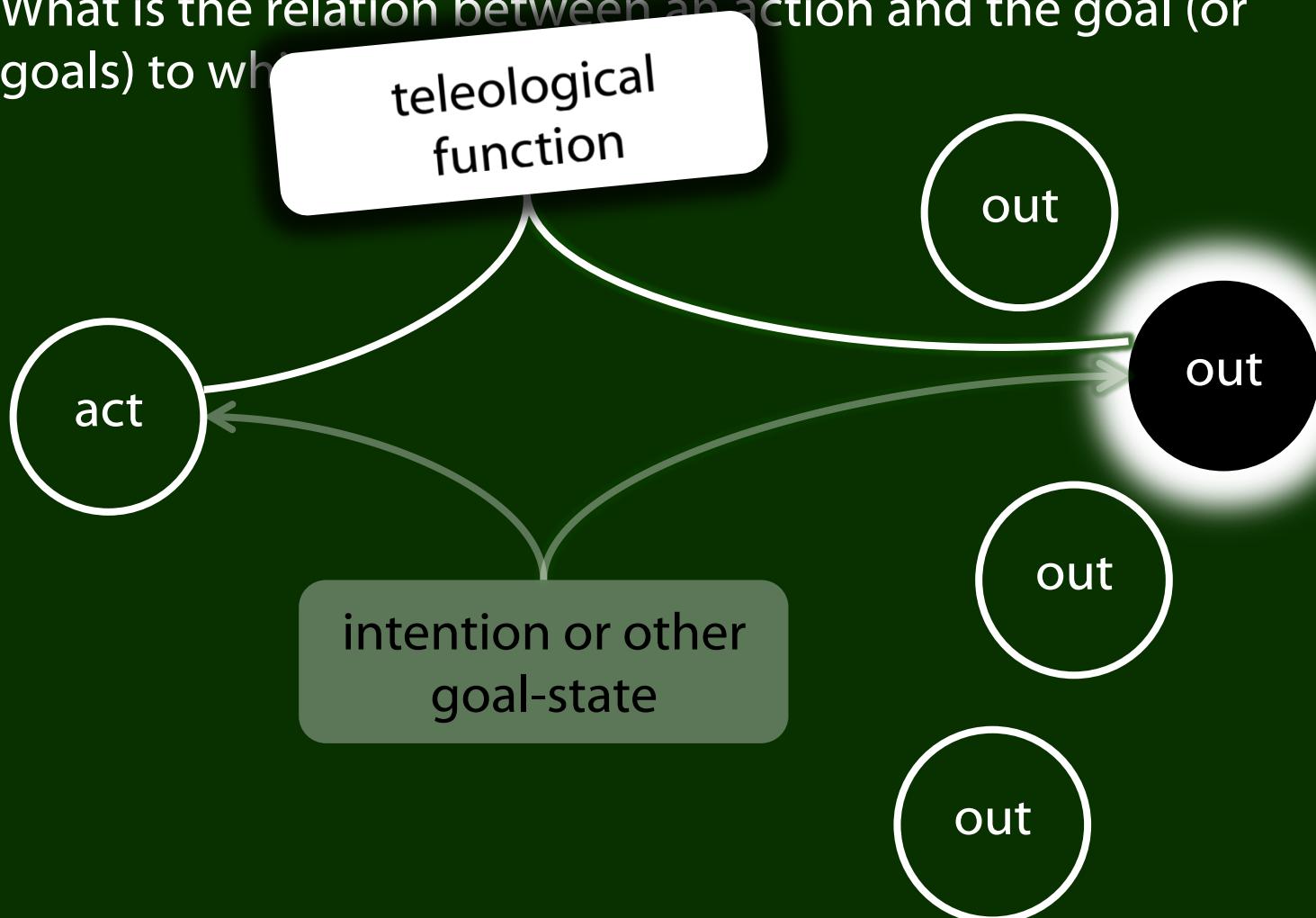
What is the relation between an action and the goal (or goals) to which it is directed?



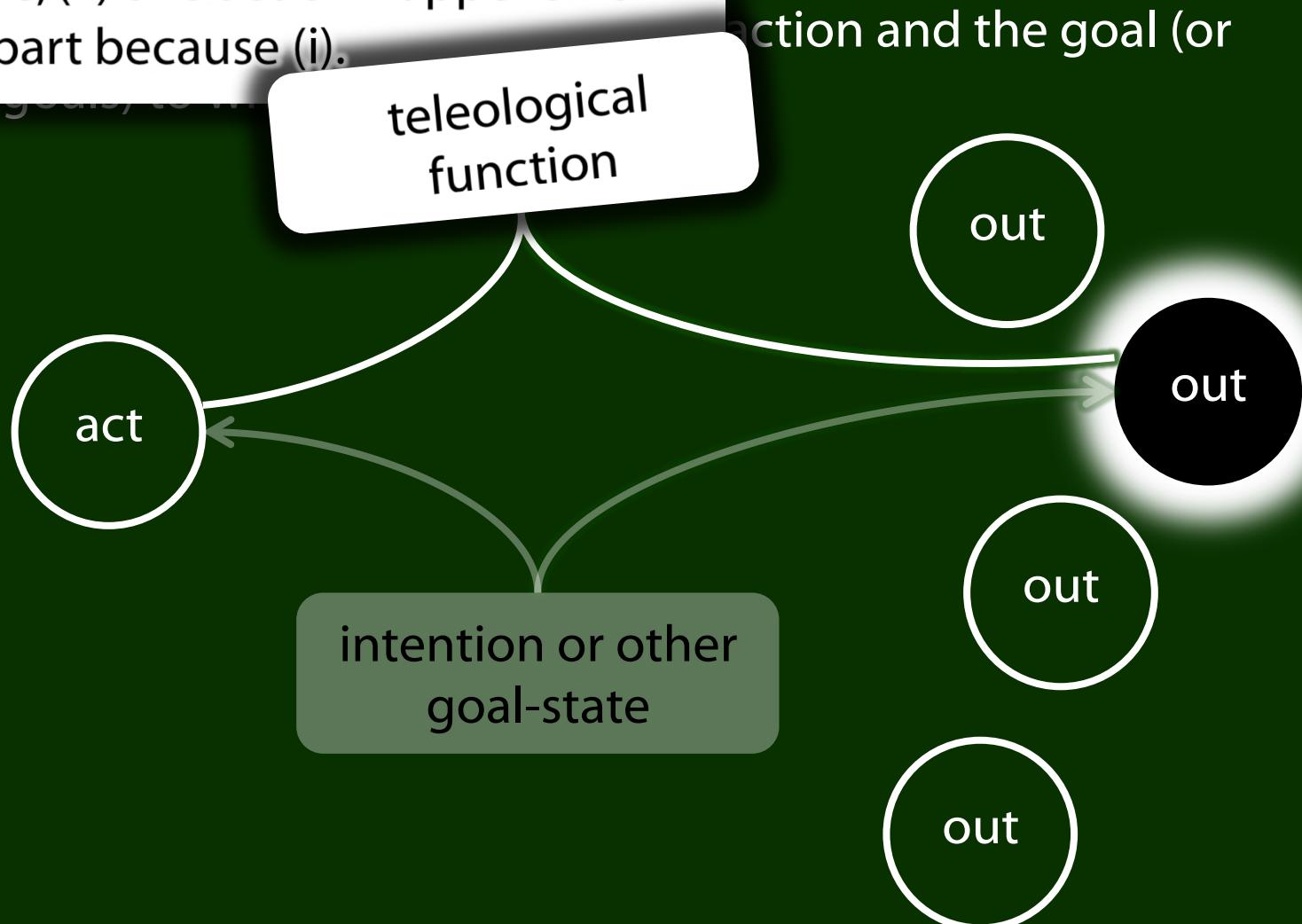
What is the relation between an action and the goal (or goals) to which it is directed?



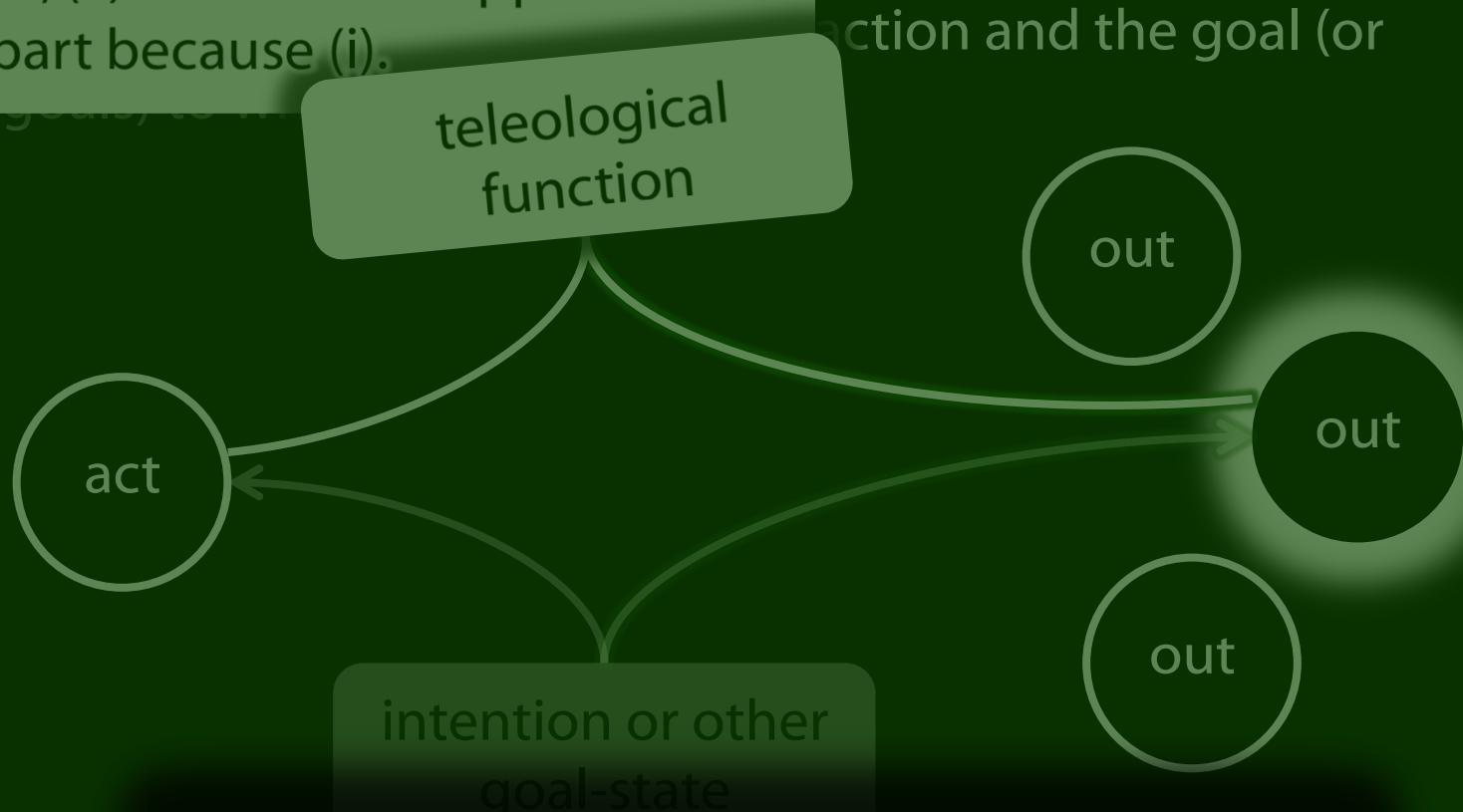
What is the relation between an action and the goal (or goals) to which it is related?



(i) in the past, actions of this type have caused outcomes of this type; (ii) this action happens now in part because (i).



(i) in the past, actions of this type have caused outcomes of this type; (ii) this action happens now in part because (i).



It is possible to represent goal-directed actions without representing intentions.

End Detour

Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

You *encounter* an object = it is in your field

Principle 3: one can't goal-directedly act on an object unless one has encountered it.



*proximity*

*orientation*

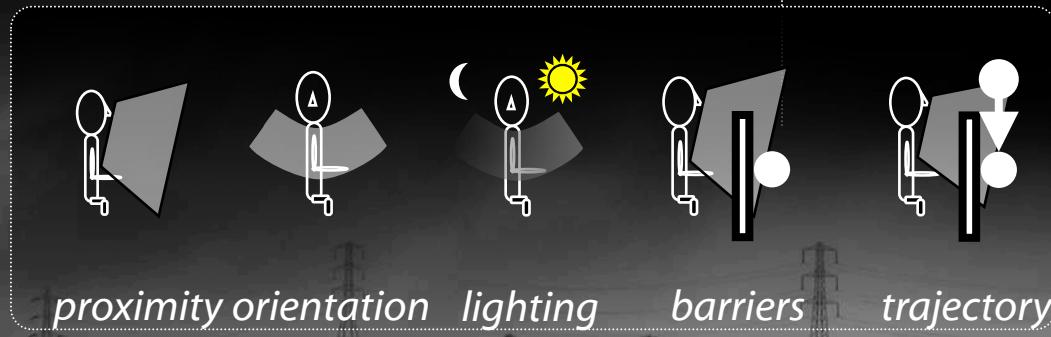
*barriers*

*trajectory*

Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

You *encounter* an object = it is in your field

Principle 3: one can't goal-directedly act on an object unless one has encountered it.

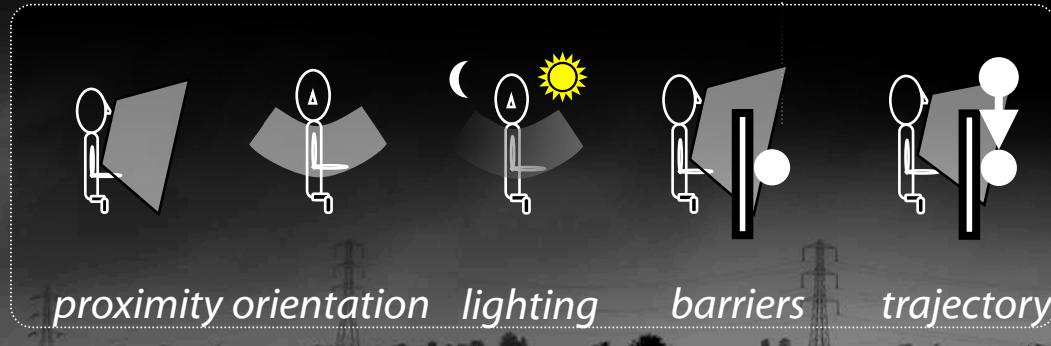


Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

“children could ... think about what the other person saw rather than what they saw”

(Flavell, Shipstead & Croft 1978: 1210)

Principle 3: one can't goal-directedly act on an object unless one has encountered it.



Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

“children could ... think about what the other person **saw** rather than what they saw”

(Flavell, Shipstead & Croft 1978: 1210)

Principle 3: one can't goal-directedly act on an object unless one has encountered it.



proximity orientation

lighting

barriers

trajectory

Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

You *encounter* an object = it is in your field

Principle 3: one can't goal-directedly act on an object unless one has encountered it.



*proximity*

*orientation*

*barriers*

*trajectory*

Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

You *encounter* an object = it is in your field

You *register* an object at a location  $\leq$  you most recently encountered the object at that location

Principle 3: one can't goal-directedly act on an object unless one has encountered it.



*proximity*

*orientation*

*lighting*

*barriers*

*trajectory*

Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

You *encounter* an object = it is in your field

You *register* an object at a location <= you most recently encountered the object at that location

Principle 3: one can't goal-directedly act on an object unless one has encountered it.

Principle 4: correct registration is a condition of *successful* action.



*proximity*

*orientation*

*barriers*

*trajectory*



source: Liszkowski et al (2008)





source: Liszkowski et al (2008)

Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

You *encounter* an object = it is in your field

You *register* an object at a location  $\leq$  you most recently encountered the object at that location



proximity



orientation



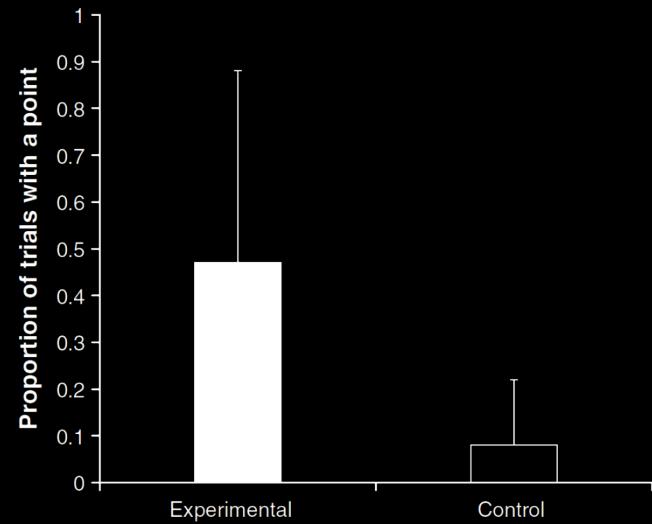
lighting



barriers

Principle 3: one can't goal-directedly act on an object unless one has encountered it.

Principle 4: correct registration is a condition of *successful* action.



source: Liszkowski et al (2008)





source: Liszkowski et al (2008)

Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

You *encounter* an object = it is in your field

You *register* an object at a location  $\leq$  you most recently encountered the object at that location



*proximity*

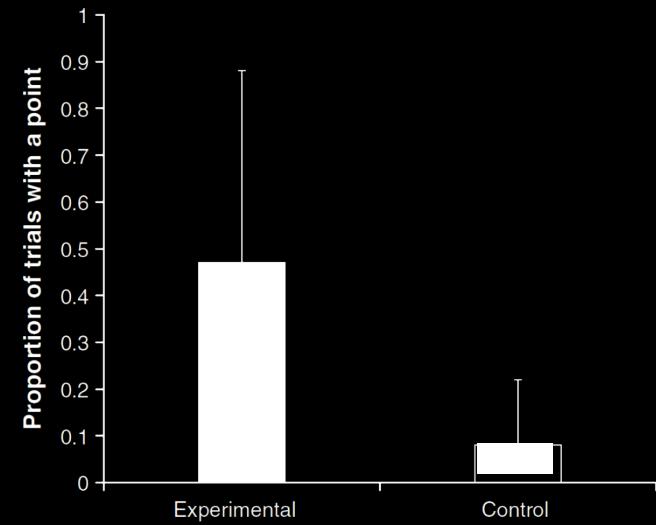
*orientation*

*barriers*

*trajectory*

Principle 3: one can't goal-directedly act on an object unless one has encountered it.

Principle 4: correct registration is a condition of *successful* action.



source: Liszkowski et al (2008)



Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

You *encounter* an object = it is in your field

“Helping by informing inextricably involves ... an understanding of others’ goals and ... of others’ ignorance.”

(Liszkowski, Carpenter & Tomasello 2008: 738-9)



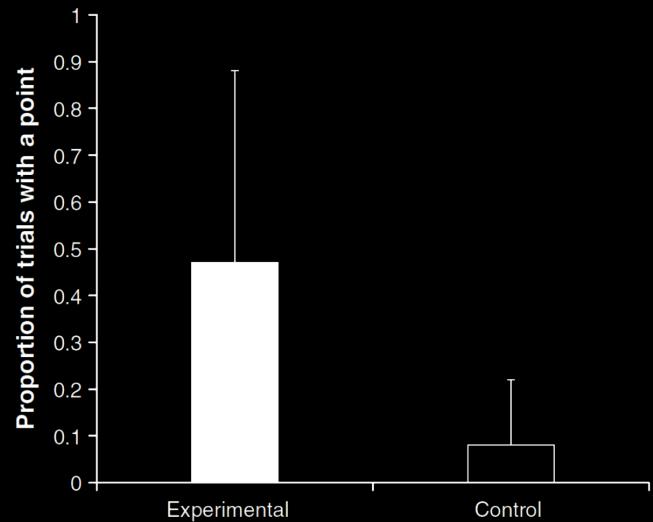
proximity orientation lighting

barriers

trajectory

Principle 3: one can’t goal-directedly act on an object unless one has encountered it.

Principle 4: correct registration is a condition of *successful* action.



source: Liszkowski et al (2008)



Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

You *encounter* an object = it is in your field

“Helping by informing inextricably involves ... an understanding of others’ goals and ... of others’ **ignorance.**”

(Liszkowski, Carpenter & Tomasello 2008: 738-9)



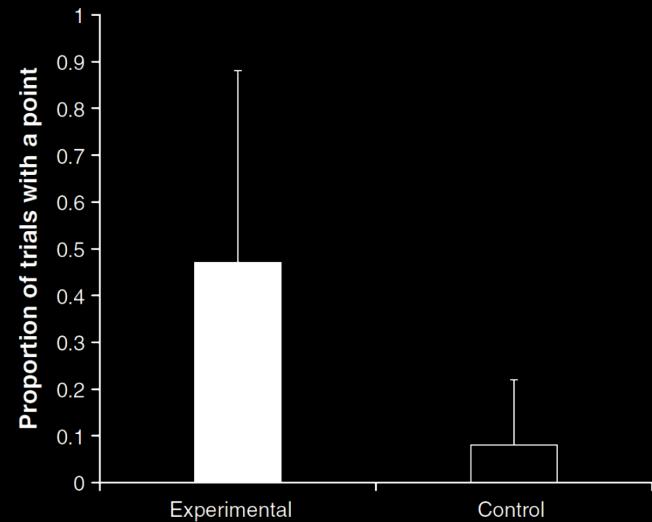
proximity orientation lighting

barriers

trajectory

Principle 3: one can’t goal-directedly act on an object unless one has encountered it.

Principle 4: correct registration is a condition of *successful* action.



source: Liszkowski et al (2008)



Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

You *encounter* an object = it is in your field

You *register* an object at a location <= you most recently encountered the object at that location

Principle 3: one can't goal-directedly act on an object unless one has encountered it.

Principle 4: correct registration is a condition of *successful* action.



*proximity*

*orientation*

*barriers*

*trajectory*

Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

You *encounter* an object = it is in your field

You *register* an object at a location <= you most recently encountered the object at that location

Principle 3: one can't goal-directedly act on an object unless one has encountered it.

Principle 4: correct registration is a condition of *successful* action.

Principle 5



*proximity*

*orientation*

*lighting*

*barriers*

*trajectory*

Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

You *encounter* an object = it is in your field

You *register* an object at a location <= you most recently encountered the object at that location

Principle 3: one can't goal-directedly act on an object unless one has encountered it.

Principle 4: correct registration is a condition of *successful* action.

Principle 5: when an agent performs a goal-directed action and the goal specifies an object, the agent will act as if the object were actually in the location she registers it at.



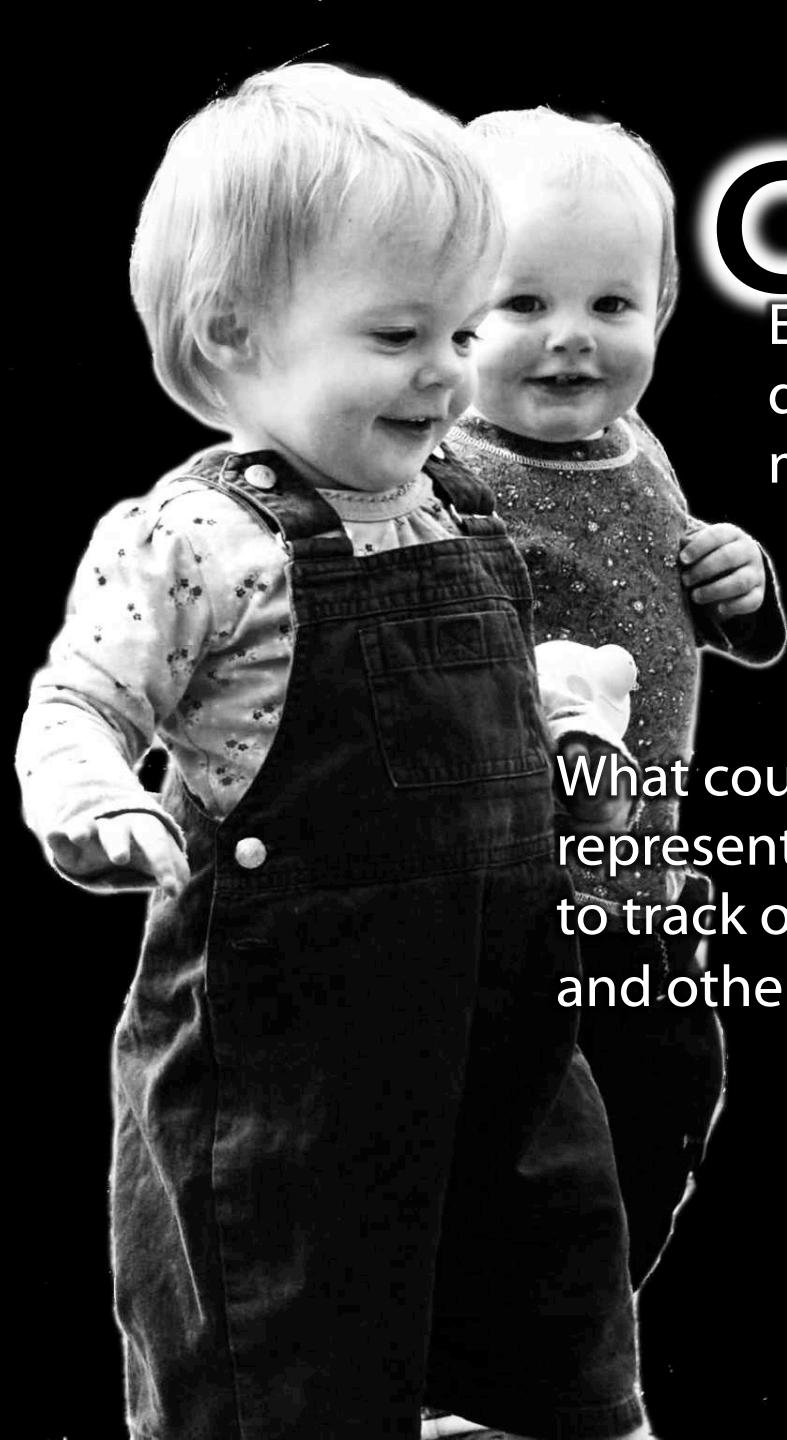
proximity

orientation

lighting

barriers

trajectory



# challenge

Explain the emergence, in evolution or development, of full-blown theory of mind cognition.

# puzzle

What could infants, chimps and scrub-jays represent that would enable them, within limits, to track others' perceptions, knowledge, beliefs and other propositional attitudes?

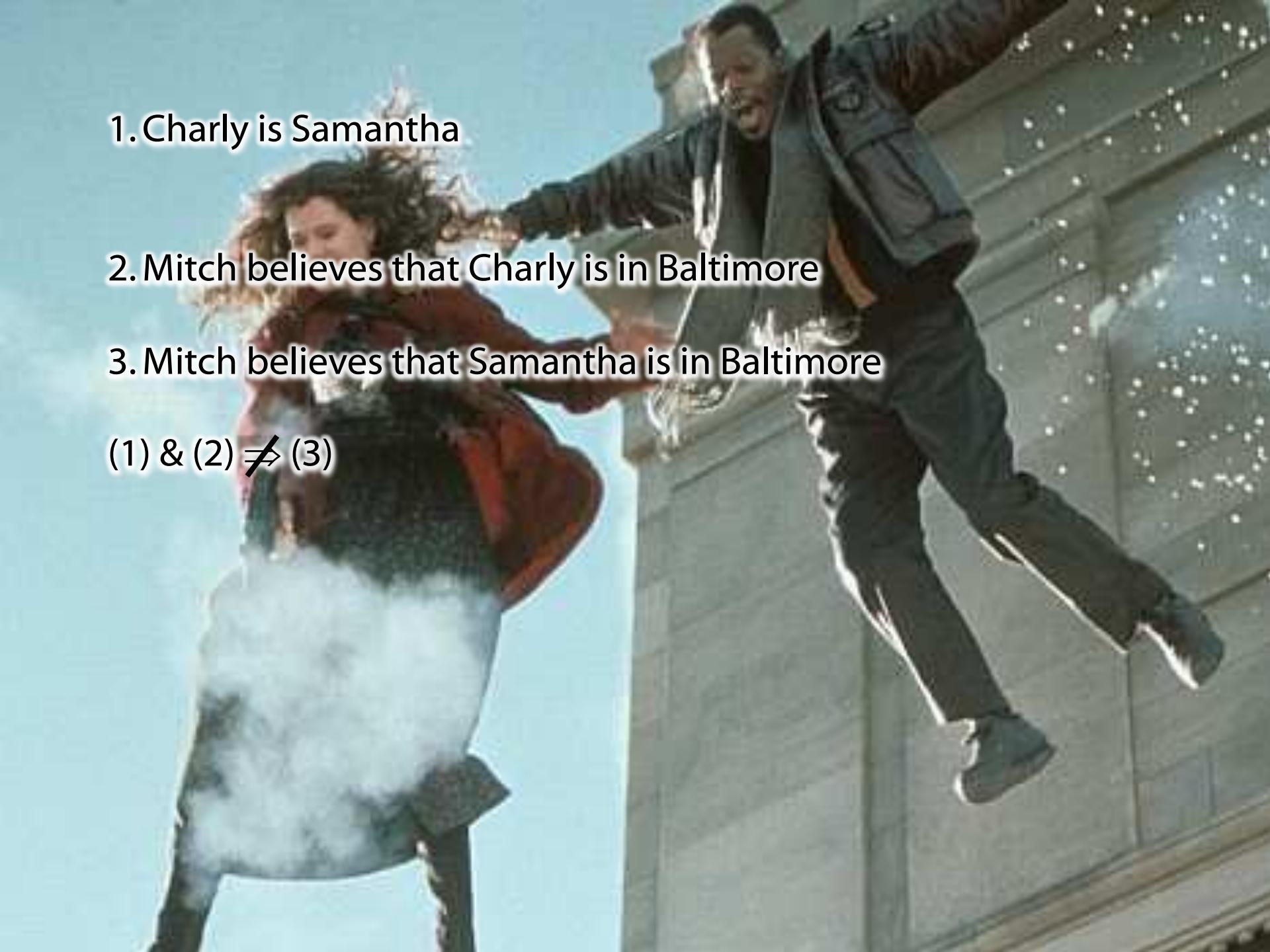
signature limits

1. Charly is Samantha

2. Mitch believes that Charly is in Baltimore

3. Mitch believes that Samantha is in Baltimore

(1) & (2)  $\not\Rightarrow$  (3)



1. Charly is Samantha

2. Mitch believes that Charly is in Baltimore

3. Mitch believes that Samantha is in Baltimore

(1) & (2)  $\not\Rightarrow$  (3)

1. Charly is Samantha
  2. Mitch believes that Charly is in Baltimore
  3. Mitch believes that Samantha is in Baltimore
- (1) & (2)  $\not\Rightarrow$  (3)
4. Mitch registers <Charly, Baltimore>
  5. Mitch registers <Samantha, Baltimore>
- (1) & (4)  $\Rightarrow$  (5)

Subjects represent  
registration only

*pass*

*fail*

Subjects represent  
beliefs

*pass*

*pass*

False belief  
about location

False belief  
about identity

Rubio, Richardson & Butterfill



Rubio, Richardson & Butterfill



Rubio, Richardson & Butterfill



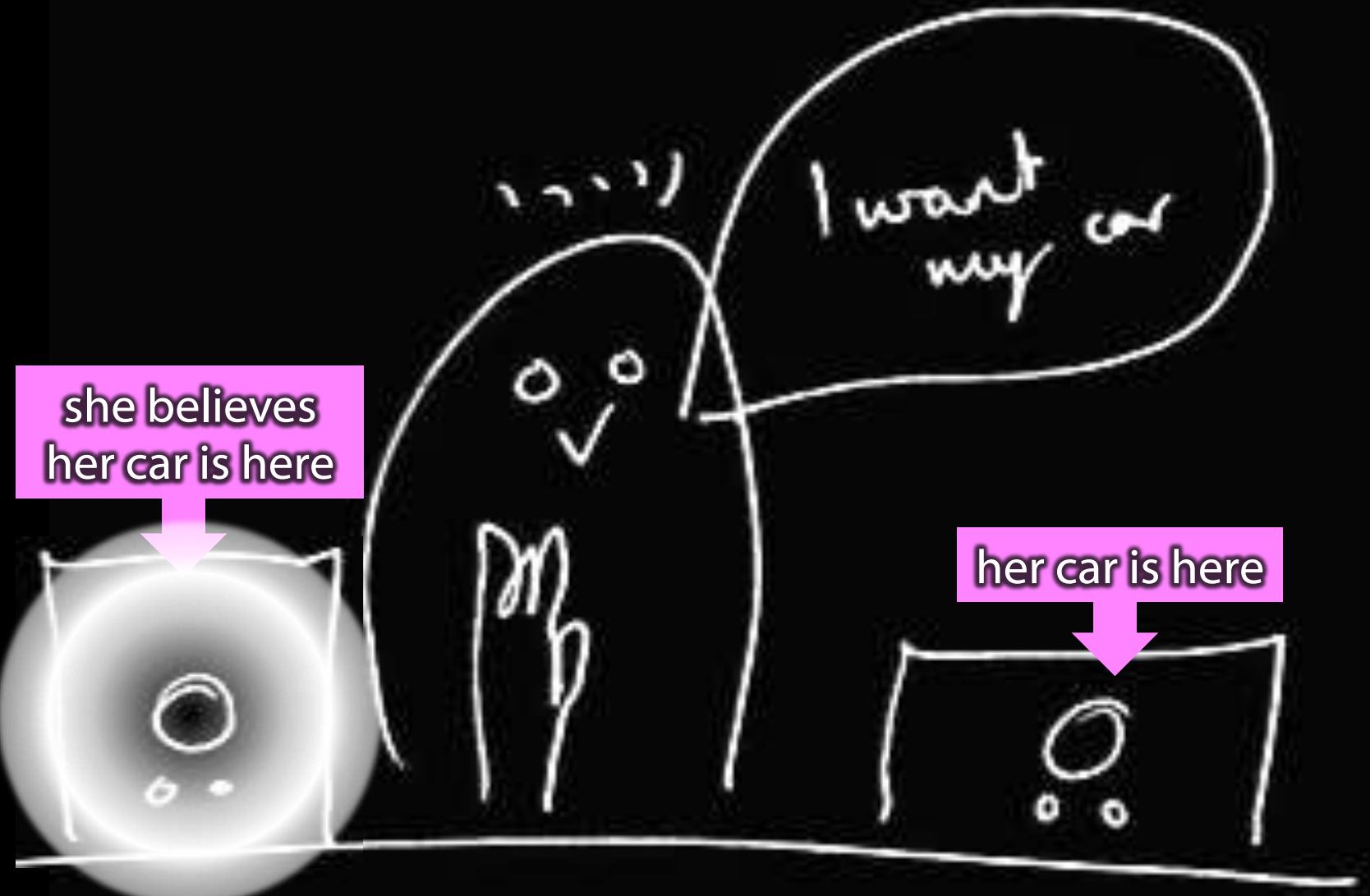
Rubio, Richardson & Butterfill



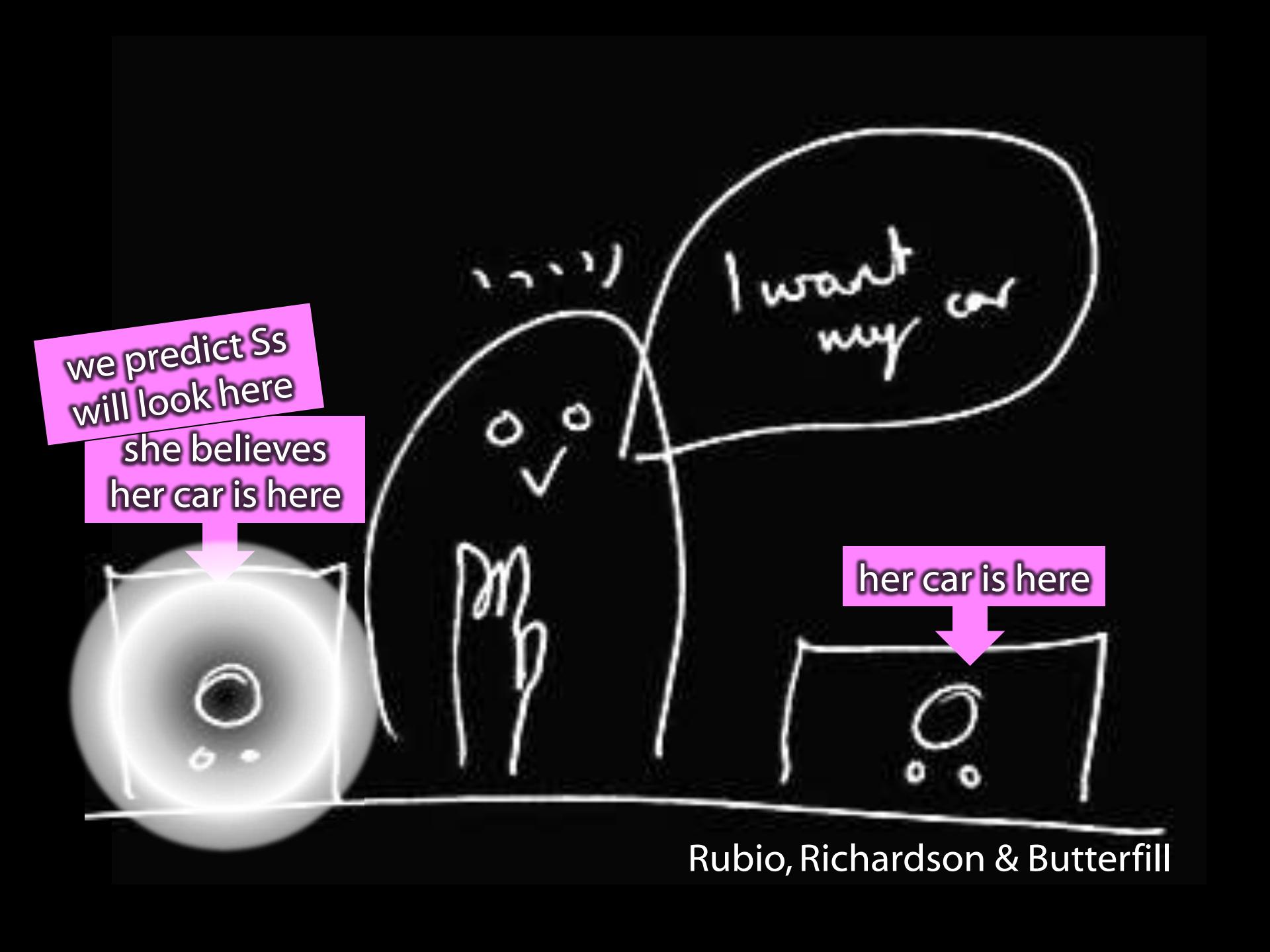
Rubio, Richardson & Butterfill



Rubio, Richardson & Butterfill



Rubio, Richardson & Butterfill



we predict Ss  
will look here

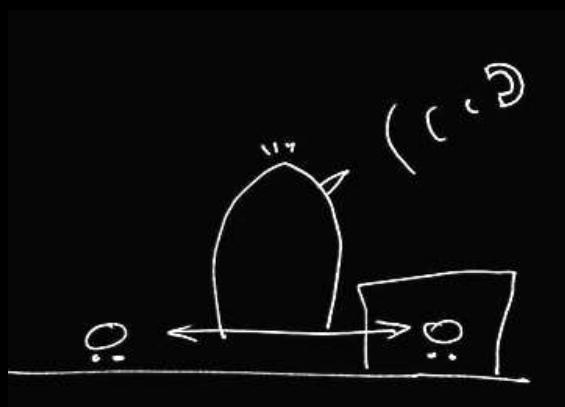
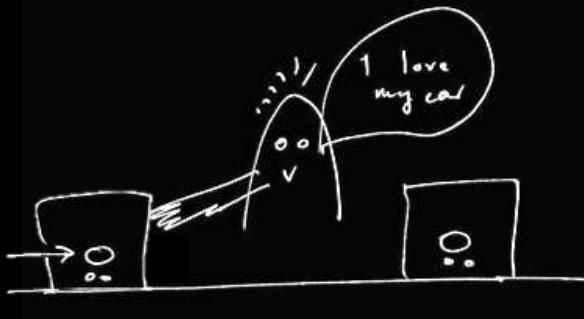
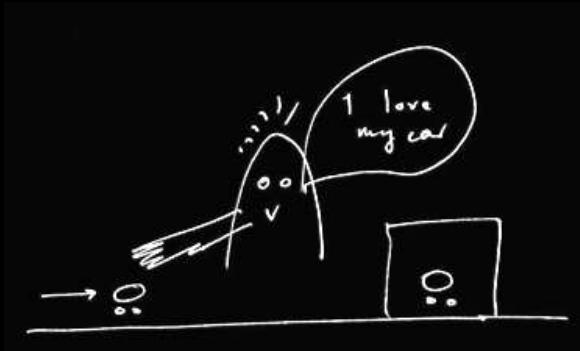
she believes  
her car is here

her car is here

Rubio, Richardson & Butterfill

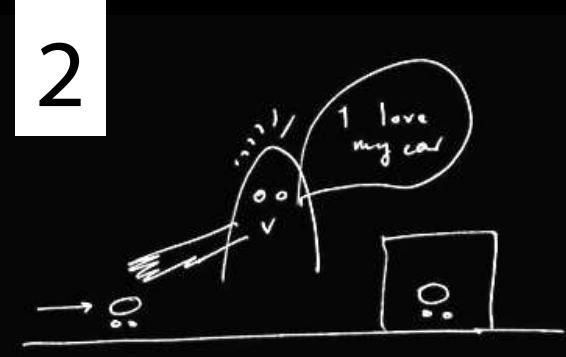


Rubio, Richardson & Butterfill

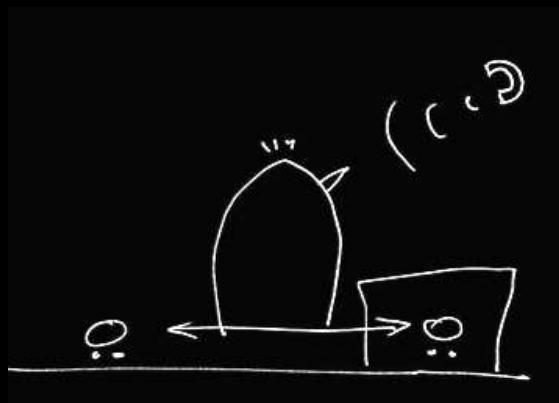
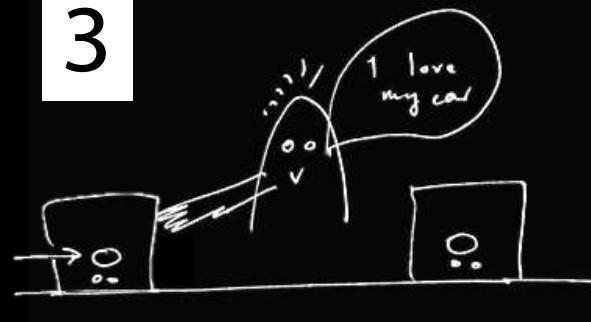


Rubio, Richardson & Butterfill

2



3



Rubio, Richardson & Butterfill

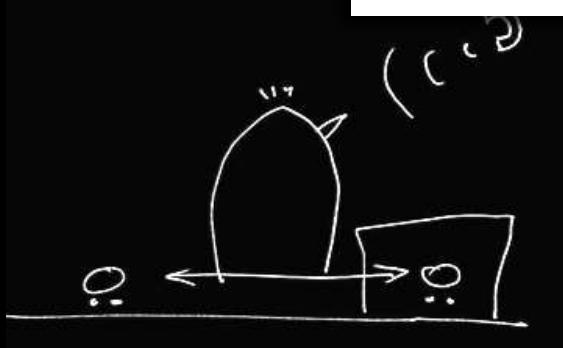
2



3



belief:  $2 \approx 3$



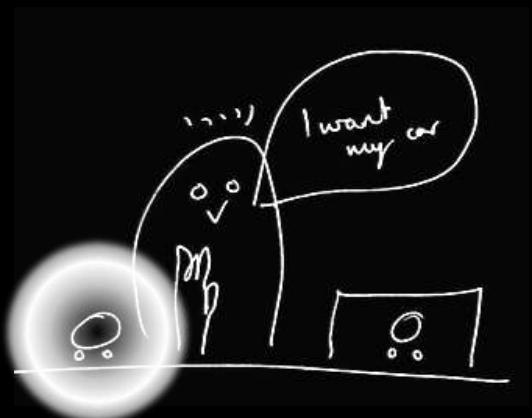
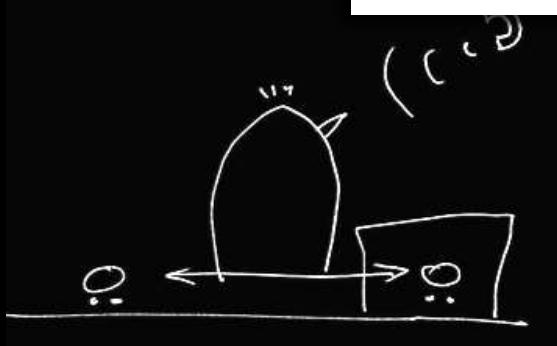
2



3



belief:  $2 \approx 3$



Rubio, Richardson & Butterfill



2



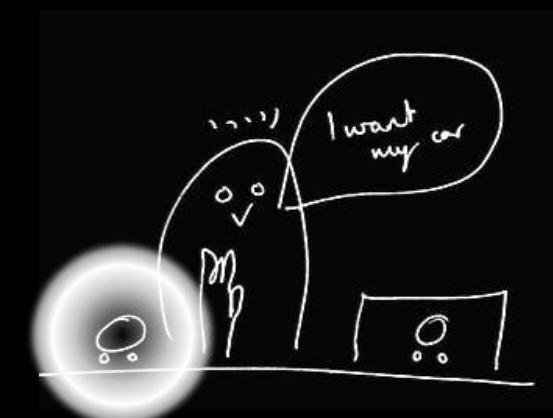
left right

left right

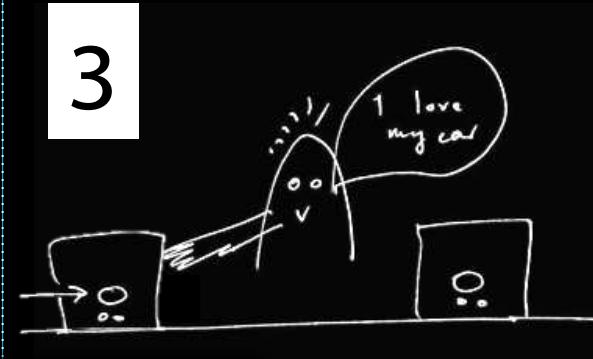
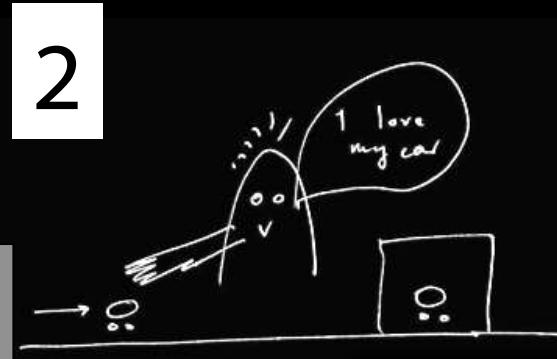
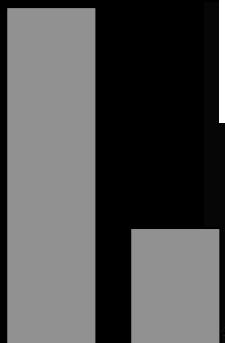
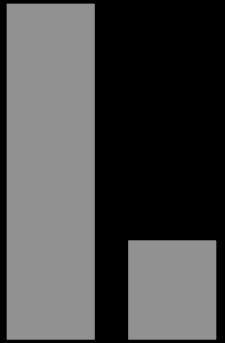
2

3

belief:  $2 \approx 3$



Rubio, Richardson & Butterfill



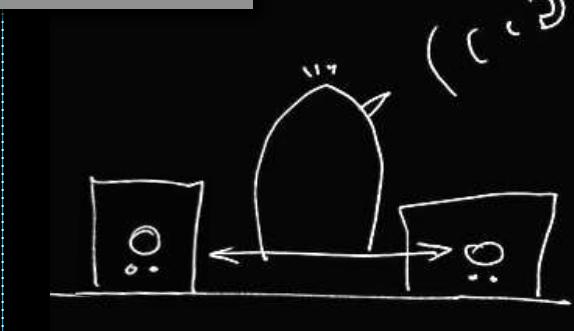
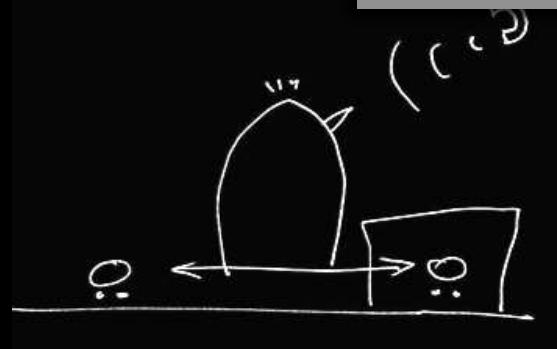
left right

2

left right

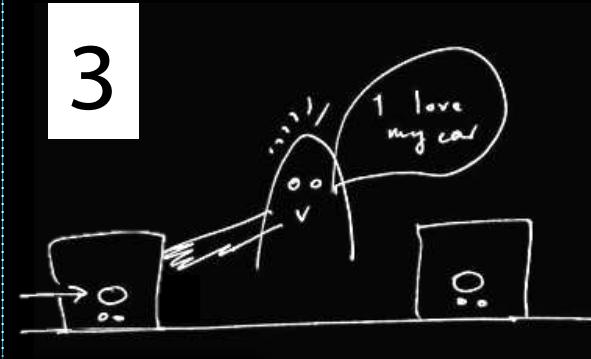
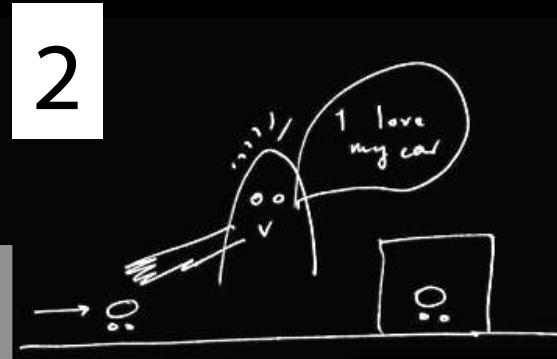
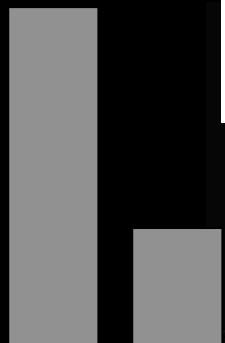
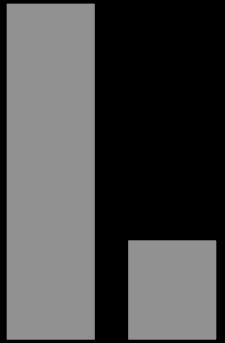
3

belief:  $2 \approx 3$



registration:  $2 \not\approx 3$





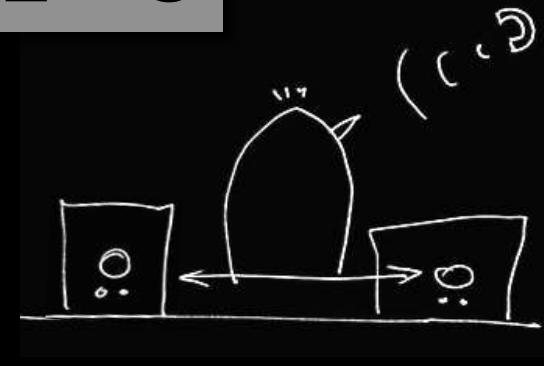
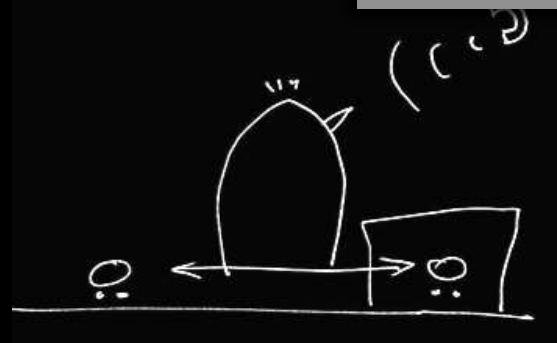
left right

2

left right

3

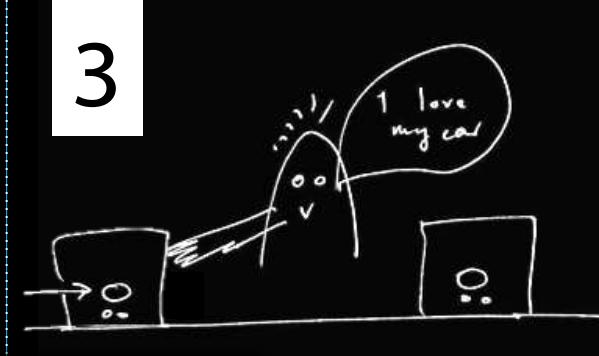
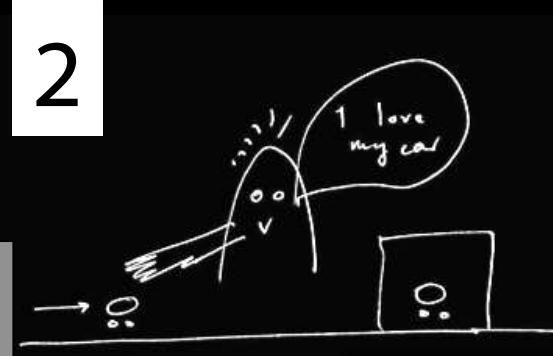
belief:  $2 \approx 3$



registration:  $2 \not\approx 3$



Rubio, Richardson & Butterfill



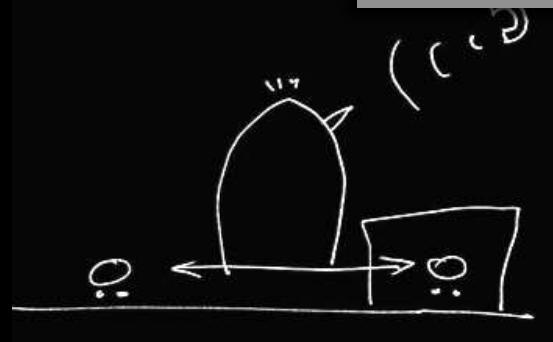
left right

2

left right

3

belief:  $2 \approx 3$



registration:  $2 \not\approx 3$



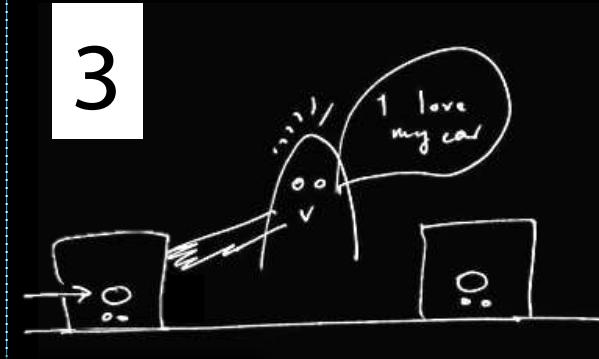
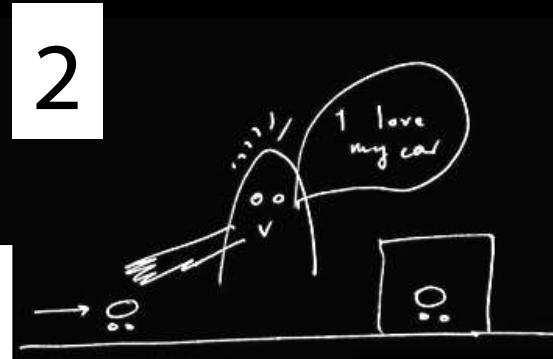
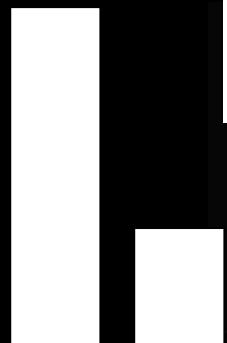
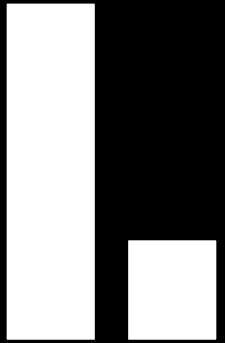
left right

2

left right

3

Rubio, Richardson & Butterfill



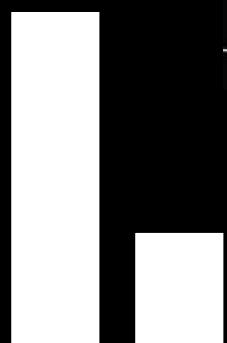
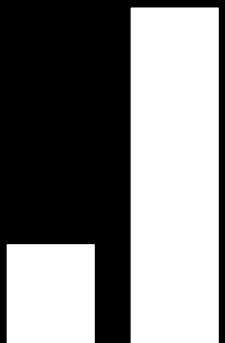
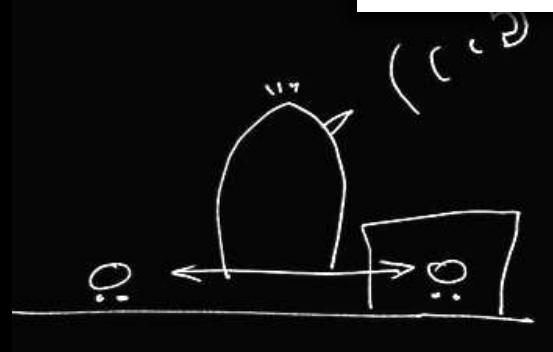
left right

left right

2

3

belief:  $2 \approx 3$



left right

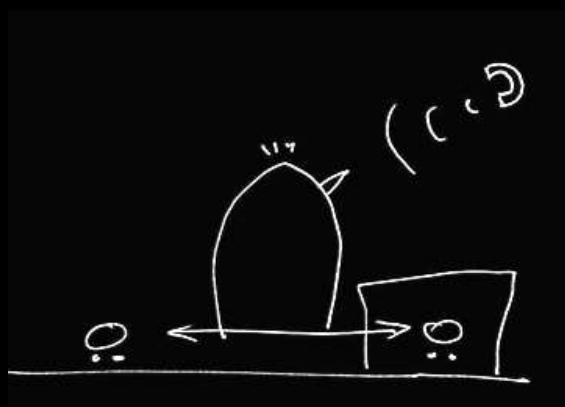
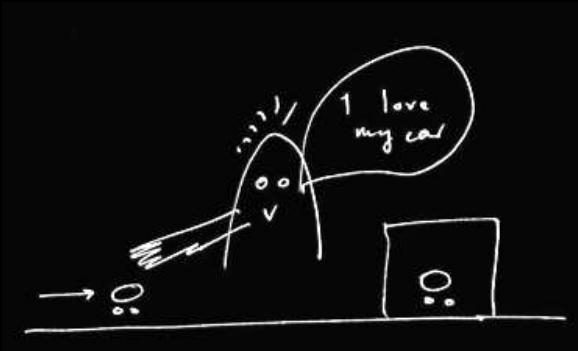
left right

2

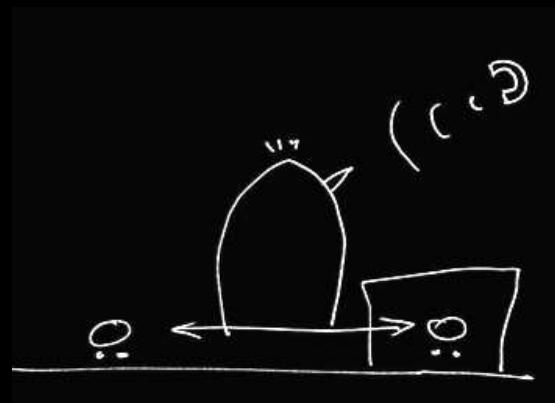
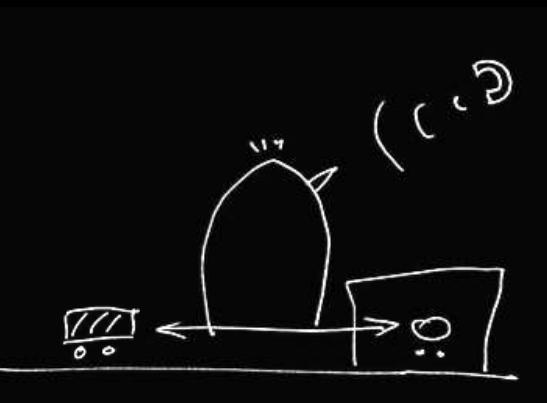
3

registration:  $2 \not\approx 3$





Rubio, Richardson & Butterfill



Rubio, Richardson & Butterfill

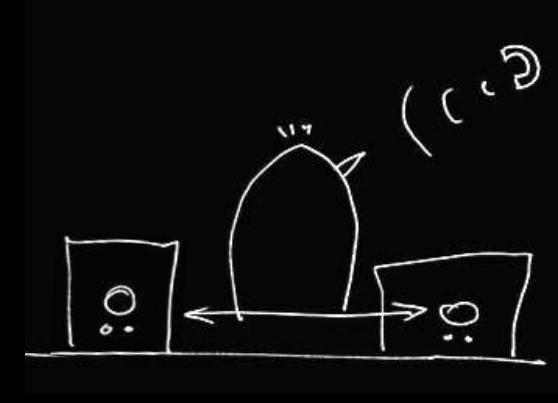
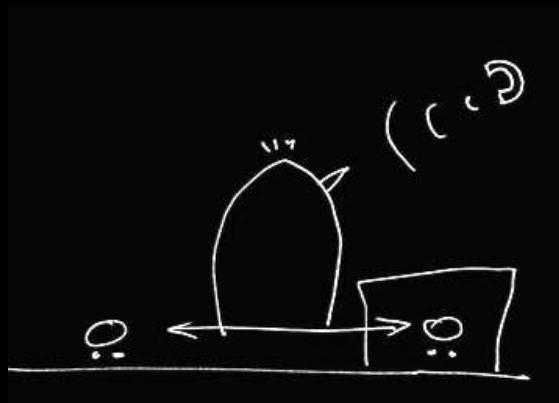
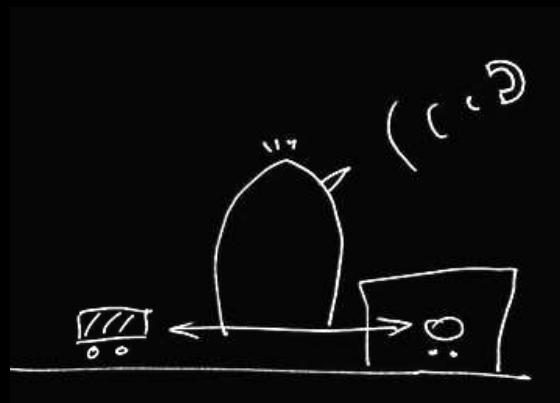
1



2



3



1



2



3



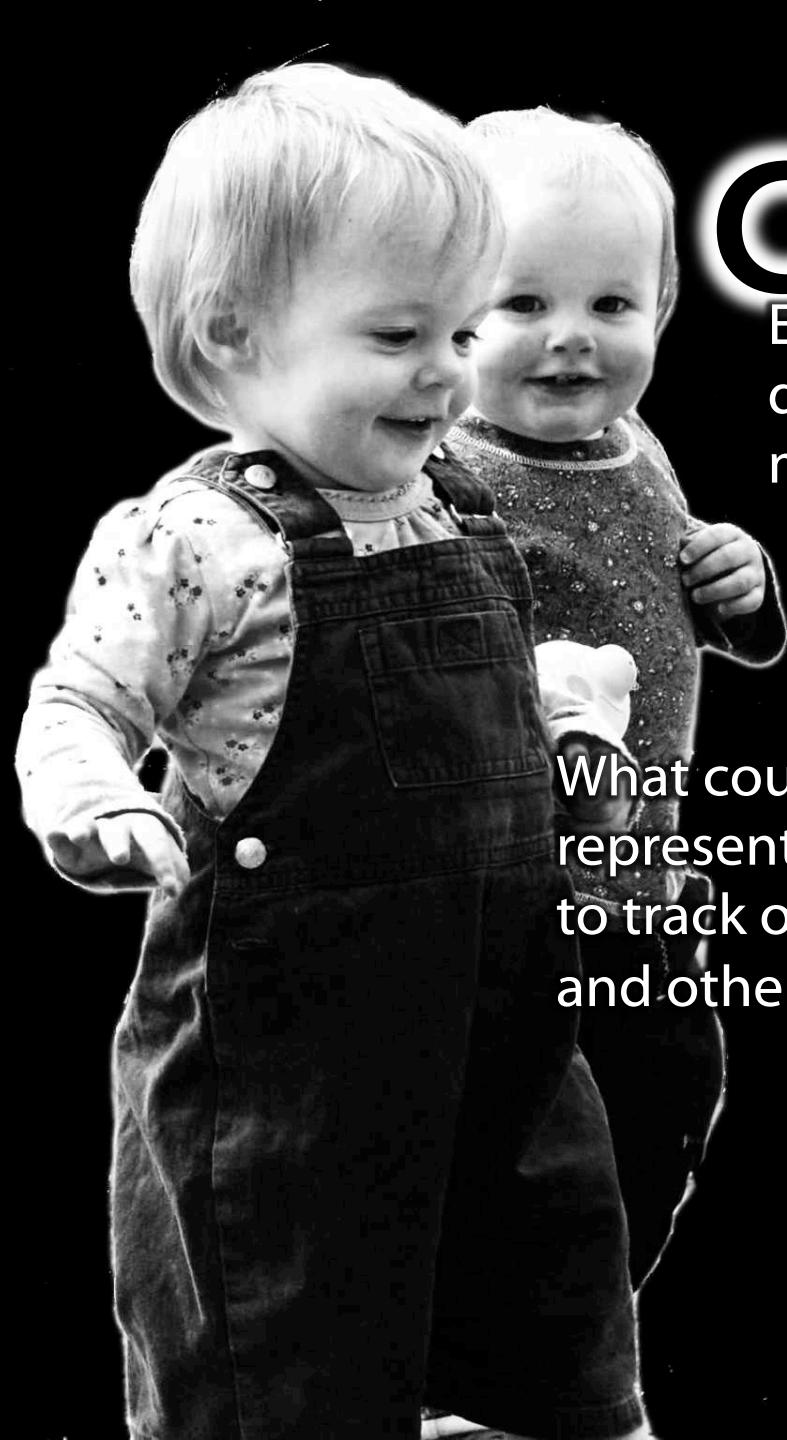
belief:  $2 \approx 3$

registration:  $2 \not\approx 3$

belief:  $1 \not\approx 2$

registration:  $1 \approx 2$





# challenge

Explain the emergence, in evolution or development, of full-blown theory of mind cognition.

# puzzle

What could infants, chimps and scrub-jays represent that would enable them, within limits, to track others' perceptions, knowledge, beliefs and other propositional attitudes?





"the unique aspects of human cognition ... were driven by, or even constituted by, social co-operation"

(Moll & Tomasello 2007)

"perception, action, and cognition are grounded in social interaction"

(Sebanz & Knoblich 2008)



# challenge

Explain the emergence, in evolution or development, of full-blown theory of mind cognition.

# puzzle

What could infants, chimps and scrub-jays represent that would enable them, within limits, to track others' perceptions, knowledge, beliefs and other propositional attitudes?

# conjecture

The existence of abilities to engage in joint action partially explains how full-blown theory of mind cognition emerges in evolution or development (or both)

A black and white photograph of two young children, a boy and a girl, smiling and holding each other. The boy is wearing overalls and a patterned shirt, and the girl is wearing a dark dress with a floral pattern. They are standing close together, suggesting a sense of connection or shared cognition.

# challenge

Explain the emergence, in evolution or development, of full-blown theory of mind cognition.

# puzzle

What could infants, chimps and scrub-jays represent that would enable them, within limits, to track others' perceptions, knowledge, beliefs and other propositional attitudes?

# conjecture

The existence of abilities to engage in **joint action** partially explains how full-blown theory of mind cognition emerges in evolution or development (or both)

tidying up the toys together

(Behne et al 2005)

cooperatively pulling  
handles in sequence to  
make a dog-puppet sing

(Brownell et al 2006)

bouncing a ball on a large  
trampoline together

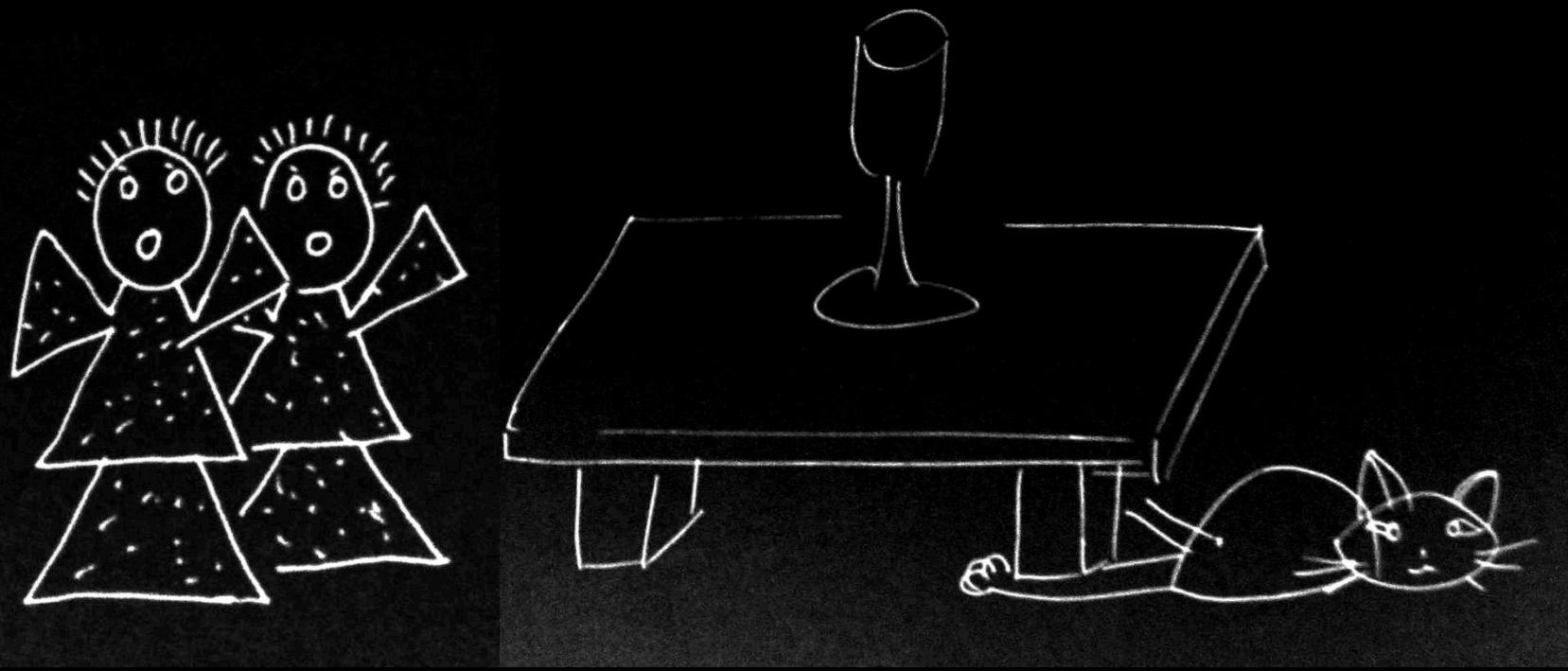
(Tomasello & Carpenter 2007)

pretending to row a boat  
together

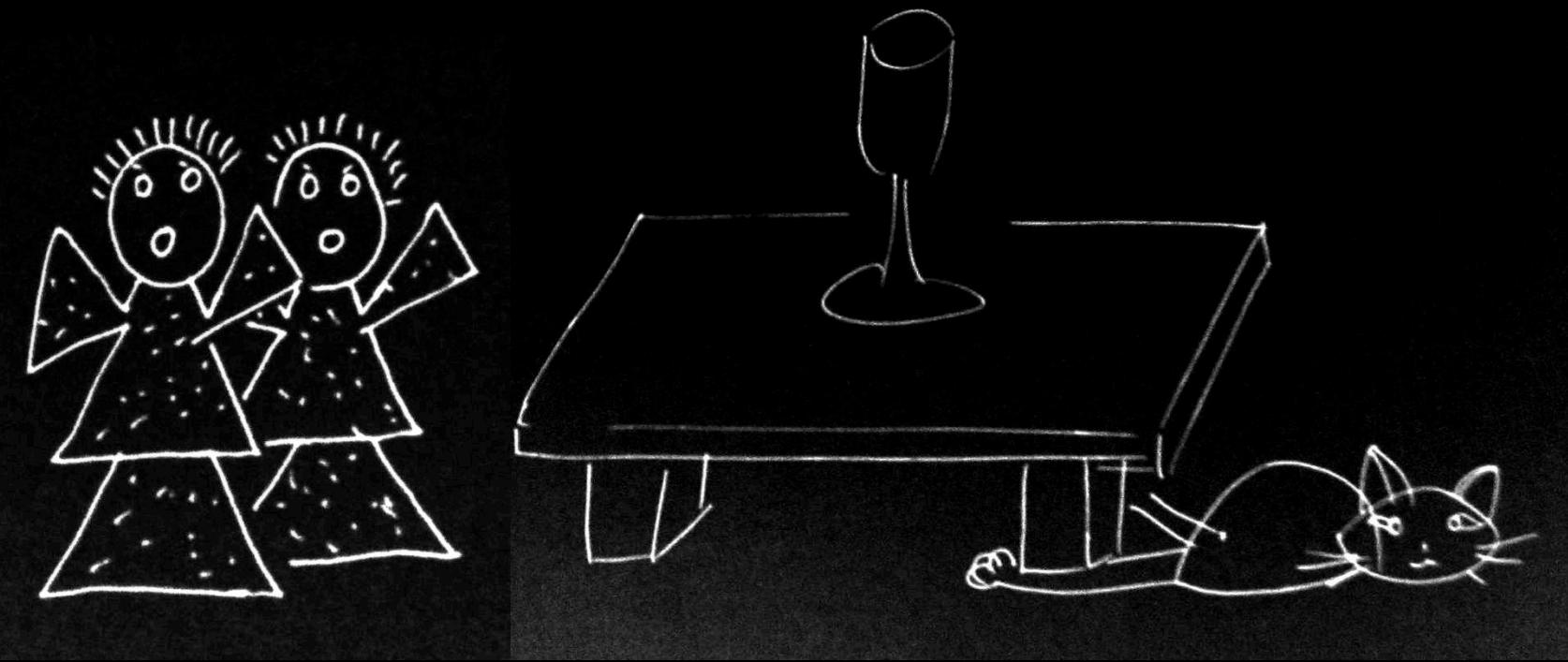
Goal-directed joint action: an event with two or more agents which, taken as a whole, is directed to a goal.

Goal-directed joint action: an event with two or more agents which, taken as a whole, is directed to a goal.

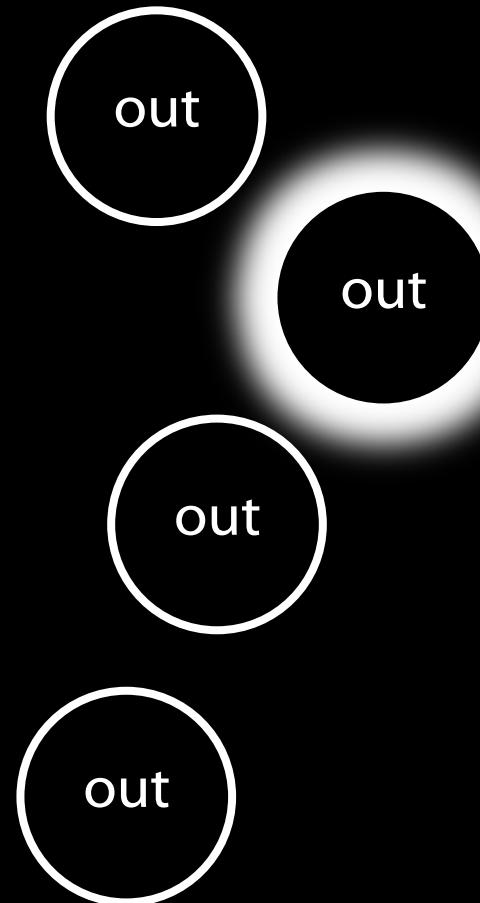
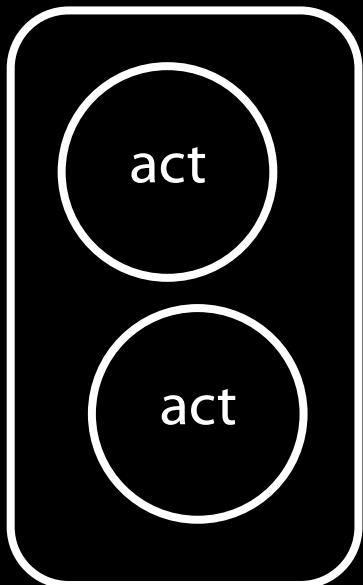
Goal-directed joint action: an event with two or more agents which, taken as a whole, is directed to a goal.



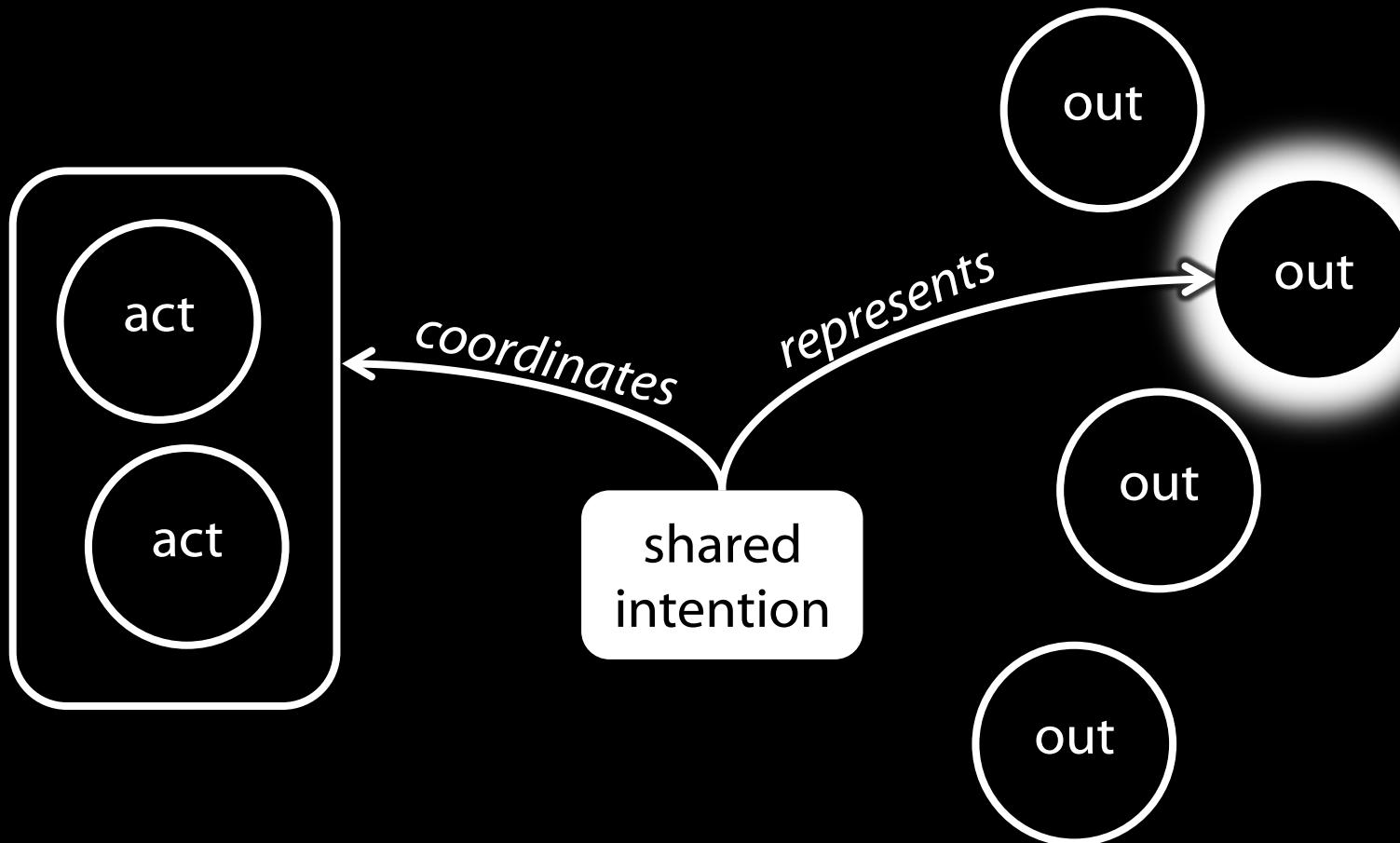
Goal-directed joint action: an event with two or more agents which, taken as a whole, is directed to a goal.



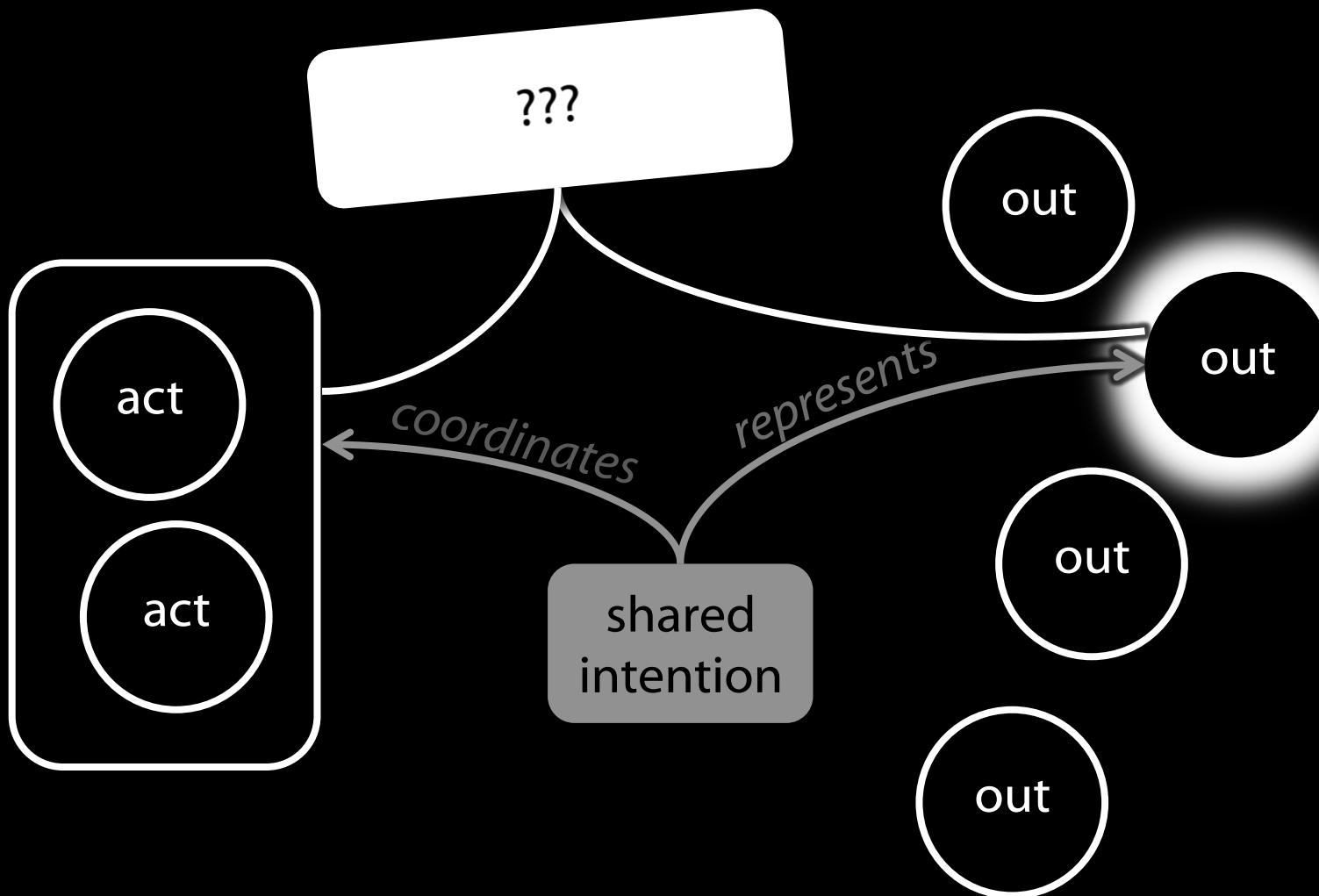
Goal-directed joint action: an event with two or more agents which, taken as a whole, is directed to a goal.



Goal-directed joint action: an event with two or more agents which, taken as a whole, is directed to a goal.



Goal-directed joint action: an event with two or more agents which, taken as a whole, is directed to a goal.

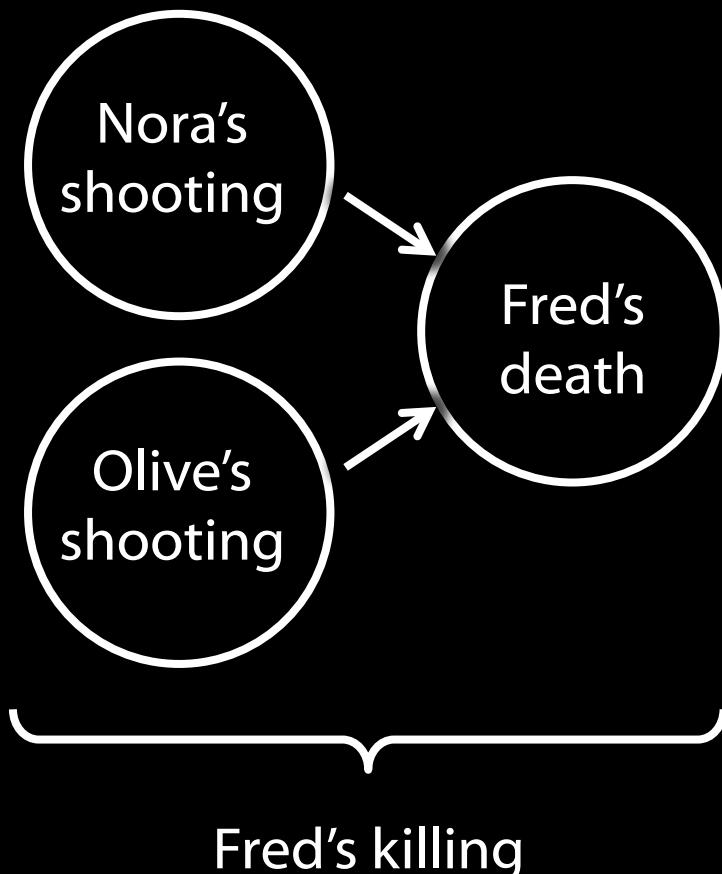


Goal-directed joint action: an event with two or more agents which, taken as a whole, is directed to a goal.

G is a distributive goal: it is an outcome to which each agent's actions are individually directed and it is possible that all actions succeed relative to this outcome.

Goal-directed joint action: an event with two or more agents which, taken as a whole, is directed to a goal.

G is a distributive goal: it is an outcome to which each agent's actions are individually directed and it is possible that: all actions succeed relative to this outcome.



Goal-directed joint action: an event with two or more agents which, taken as a whole, is directed to a goal.

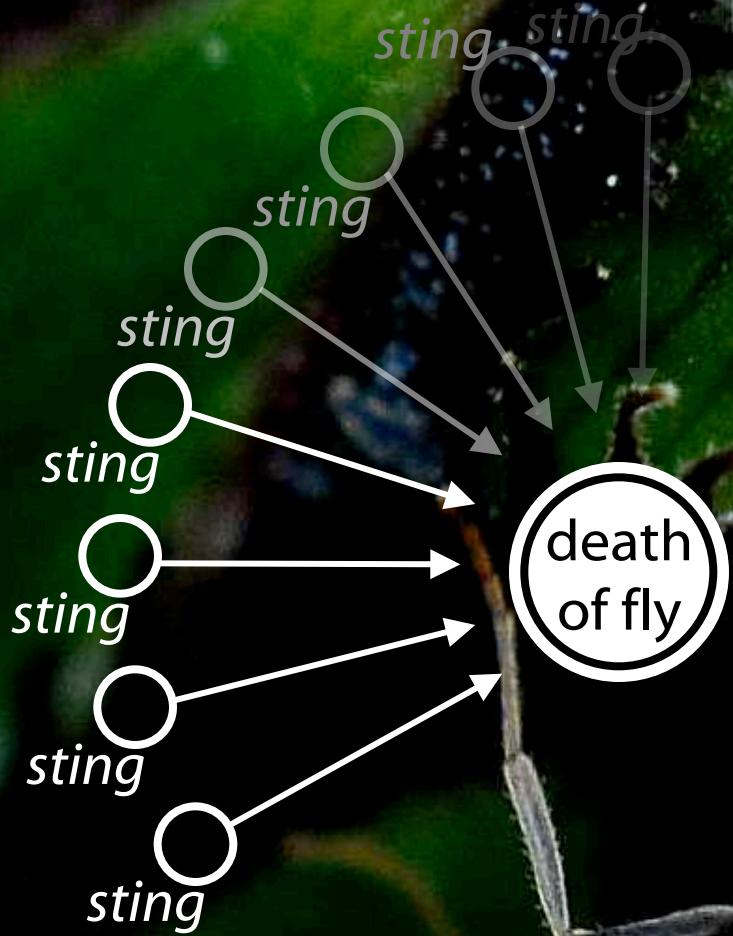
G is a distributive goal: it is an outcome to which each agent's actions are individually directed and it is possible that all actions succeed relative to this outcome.

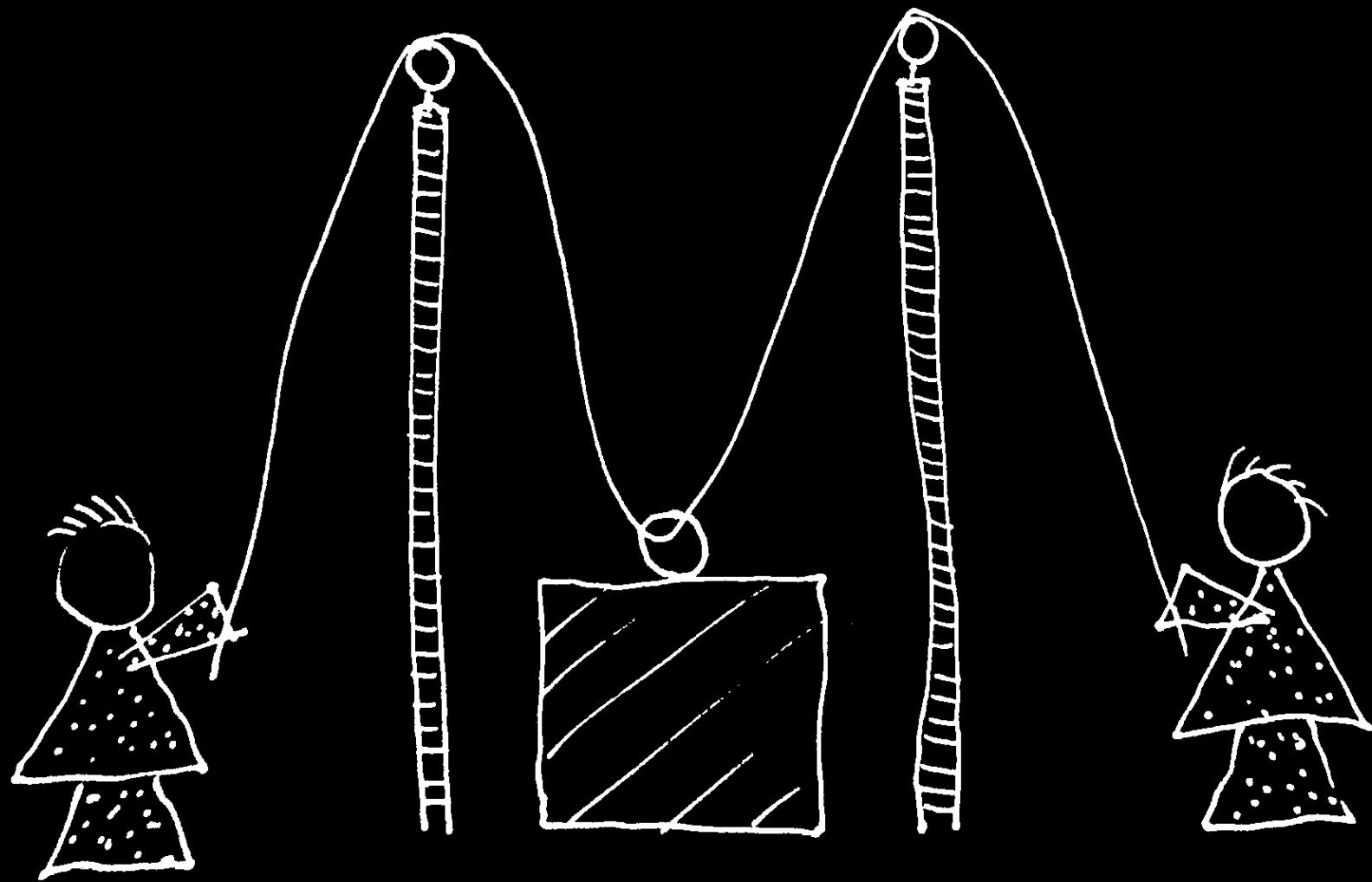
G is a collective goal

---

- (a) it is a distributive goal;
- (b) the actions are coordinated; and
- (c) coordination of this type would normally facilitate occurrences of outcomes of this type.







Goal-directed joint action: an event with two or more agents which, taken as a whole, is directed to a goal.

G is a distributive goal: it is an outcome to which each agent's actions are individually directed and it is possible that all actions succeed relative to this outcome.

G is a collective goal

- (a) it is a distributive goal;
- (b) the actions are coordinated; and
- (c) coordination of this type would normally facilitate occurrences of outcomes of this type.

Goal-directed joint action: an event with two or more agents which, taken as a whole, is directed to a goal.

G is a distributive goal: it is an outcome to which each agent's actions are individually directed and it is possible that: all actions succeed relative to this outcome.

G is a shared goal

---

G is a collective goal

- (a) it is a distributive goal;
- (b) the actions are coordinated; and
- (c) coordination of this type would normally facilitate occurrences of outcomes of this type.

Each agent most wants and expects each of the other agents to perform activities directed to the goal.

Each agent most wants and expects the goal to occur as a common effect of all their goal-directed actions.

Goal-directed joint action: an event with two or more agents which, taken as a whole, is directed to a goal.

G is a distributive goal: it is an outcome to which each agent's actions are individually directed and it is possible that: all actions succeed relative to this outcome.

G is a shared goal

---

G is a collective goal

- (a) it is a distributive goal;
- (b) the actions are coordinated; and
- (c) coordination of this type would normally facilitate occurrences of outcomes of this type.

Each agent **most wants** and expects each of the other agents to perform activities directed to the goal.

Each agent most wants and expects the goal to occur as a common effect of all their goal-directed actions.

Goal-directed joint action: an event with two or more agents which, taken as a whole, is directed to a goal.

G is a distributive goal: it is an outcome to which each agent's actions are individually directed and it is possible that: all actions succeed relative to this outcome.

G is a shared goal

---

G is a collective goal

- (a) it is a distributive goal;
- (b) the actions are coordinated; and
- (c) coordination of this type would normally facilitate occurrences of outcomes of this type.

Each agent most wants and expects each of the other agents to perform activities directed to the goal.

Each agent most wants and expects the goal to occur as a common effect of all their goal-directed actions.

Goal-directed joint action: an event with two or more agents which, taken as a whole, is directed to a goal.

G is a distributive goal: it is an outcome to which each agent's actions are individually directed and it is possible that: all actions succeed relative to this outcome.

G is a shared goal

---

G is a collective goal

- (a) it is a distributive goal;
- (b) the actions are coordinated; and
- (c) coordination of this type would normally facilitate occurrences of outcomes of this type.

Each agent most wants and expects each of the other agents to perform activities directed to the goal.

Each agent most wants and expects the goal to occur as a common effect of all their goal-directed actions.

Goal-directed joint action: an event with two or more agents which, taken as a whole, is directed to a goal.

G is a distributive goal: it is an outcome to which each agent's actions are individually directed and it is possible that all actions succeed relative to this outcome.

etc ...

---

### G is a shared goal

---

---

#### G is a collective goal

---

- (a) it is a distributive goal;
- (b) the actions are coordinated; and
- (c) coordination of this type would normally facilitate occurrences of outcomes of this type.

Each agent most wants and expects each of the other agents to perform activities directed to the goal.

Each agent most wants and expects the goal to occur as a common effect of all their goal-directed actions.



# challenge

Explain the emergence, in evolution or development, of full-blown theory of mind cognition.

# puzzle

What could infants, chimps and scrub-jays represent that would enable them, within limits, to track others' perceptions, knowledge, beliefs and other propositional attitudes?

# conjecture

The existence of abilities to engage in joint action partially explains how full-blown theory of mind cognition emerges in evolution or development (or both)

failed reach



point



source: Hare & Tomasello (2004)

The problem of opaque means

**your-goal-is-my-goal**

## **your-goal-is-my-goal**

1. We are about to engage in some joint action\* or other
2. I am not about to change my goal.

Therefore:

3. Your actions also will be directed to this goal.

[\*in at least the minimal sense associated with distributive goals]

failed reach



point



source: Hare & Tomasello (2004)

failed reach



point



source: Hare & Tomasello (2004)

“to understand pointing, the subject needs to understand more than the individual goal-directed behaviour. She needs to understand that ... the other attempts to communicate to her ... and ... the communicative intention behind the gesture”

(Moll & Tomasello 2007)

failed reach



point



source: Hare & Tomasello (2004)

“to understand pointing, the subject needs to understand more than the individual goal-directed behaviour. She needs to understand that ... the other attempts to communicate to her ... and ... the **communicative intention** behind the gesture”

(Moll & Tomsello 2007)



# challenge

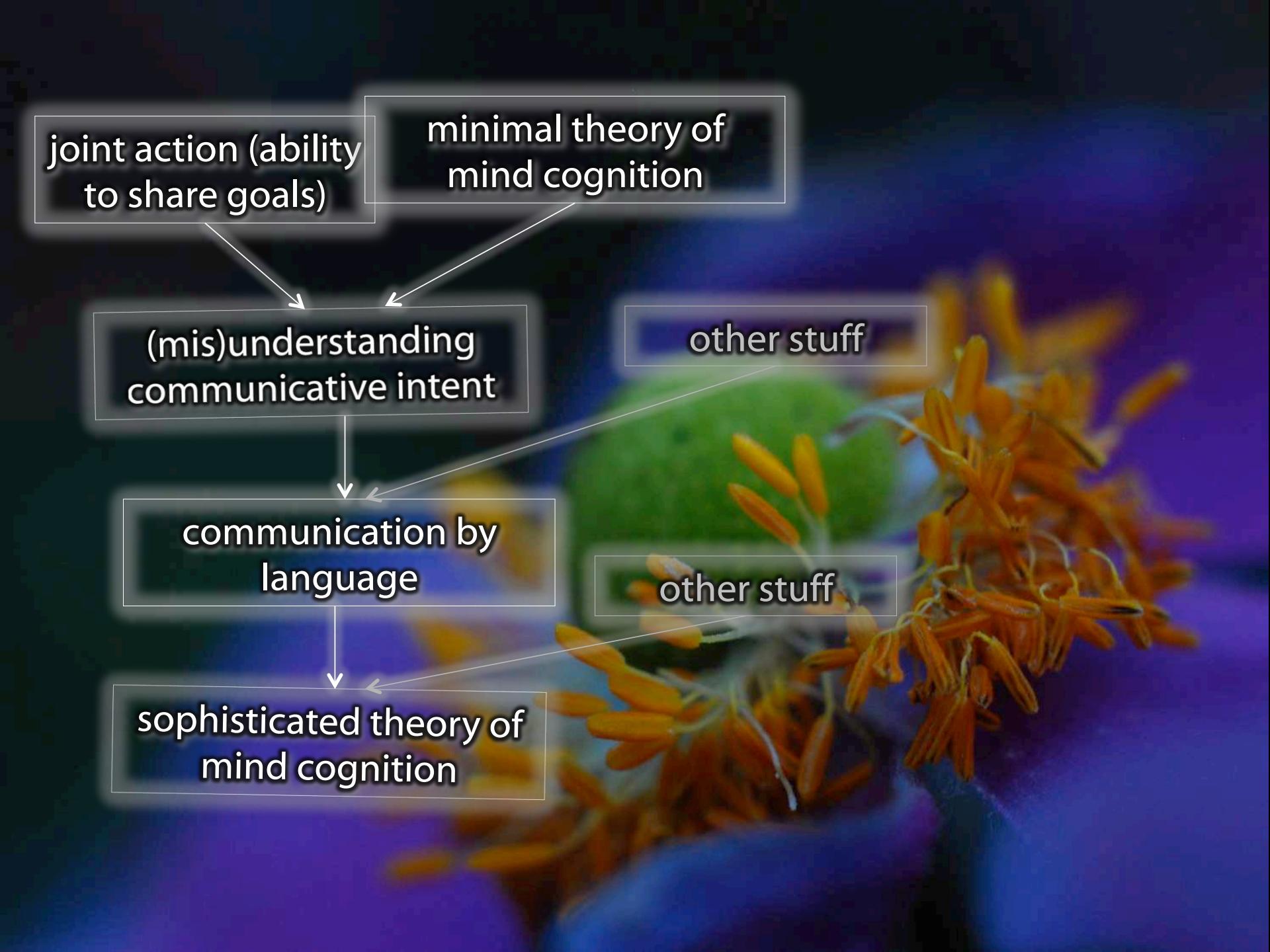
Explain the emergence, in evolution or development, of full-blown theory of mind cognition.

# puzzle

What could infants, chimps and scrub-jays represent that would enable them, within limits, to track others' perceptions, knowledge, beliefs and other propositional attitudes?

# conjecture

The existence of abilities to engage in joint action partially explains how full-blown theory of mind cognition emerges in evolution or development (or both)



joint action (ability  
to share goals)

minimal theory of  
mind cognition

(mis)understanding  
communicative intent

communication by  
language

sophisticated theory of  
mind cognition

other stuff

other stuff

