Monitoring and controlling the mental states of others

Stephen A. Butterfill & Ian A. Apperly

Mindreading makes contradictory demands

Apperly & Butterfill (2009) *Psych. Rev.*

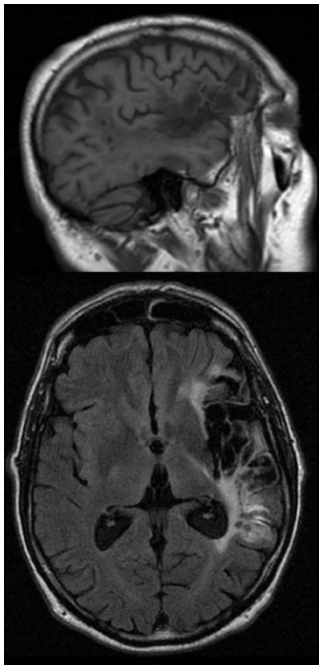
ToM must be flexible
- An archetypal "central process"

ToM must be fast and efficient
- An archetypal "modular process"

Fast &
Flexible?

Evidence that mindreading is a flexible but demanding ability

- **In Adults....**
- Impaired executive processes can lead to severe egocentrism
  - (e.g., Samson, Apperly, Kathirgamanathan & Humphreys, 2005)



Reality-unknown FB task:

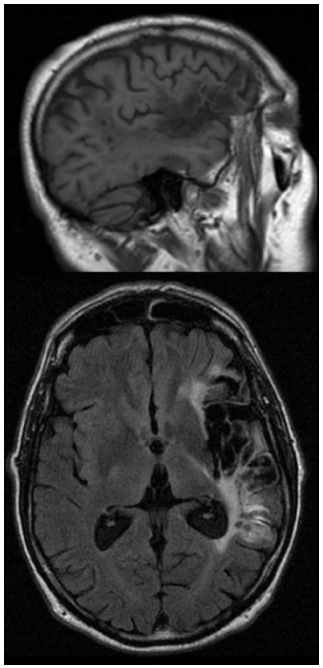
LOW SELF-perspective inhibition

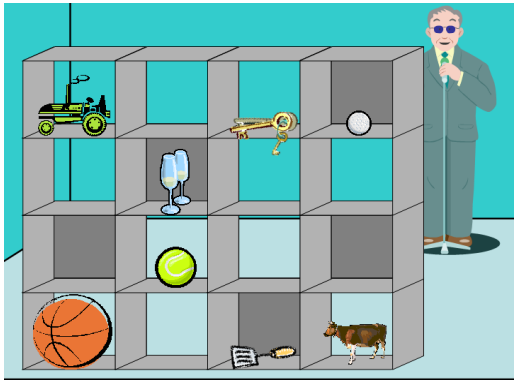
The participant does NOT know where the object is

Inferring the woman's false belief does NOT require SELF-perspective inhibition

Evidence that mindreading is a flexible but demanding ability

- **In Adults....**
- Impaired executive processes can lead to severe egocentrism
  - (e.g., Samson, Apperly, Kathirgamanathan & Humphreys, 2005)
- Belief reasoning requires cognitive control
  - (e.g., Bull, Philips & Conway, 2007)
- Belief inferences are not *made* automatically
  - (Apperly, Samson, Riggs, Simpson & Chiavarino, 2006; Back & Apperly, 2010)
- Belief inferences are not *used* automatically
  - (e.g., Keysar, Lin & Barr, 2003; Apperly et al., 2010)
- Holding false beliefs briefly in mind has a measurable processing cost
  - (Apperly, Back et al., 2008)
- Recursion (e.g., beliefs about beliefs) remains challenging
  - E.g., Mckinnon & Moscovitch (2007)
- **And of course in children...**



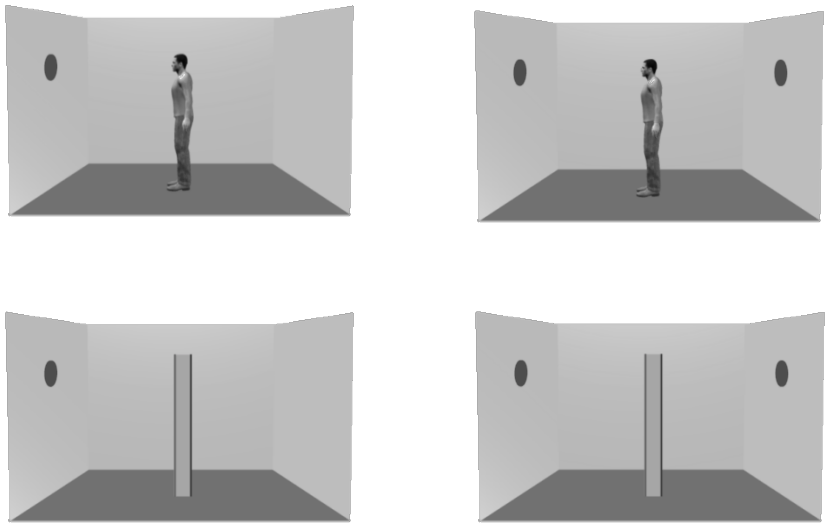


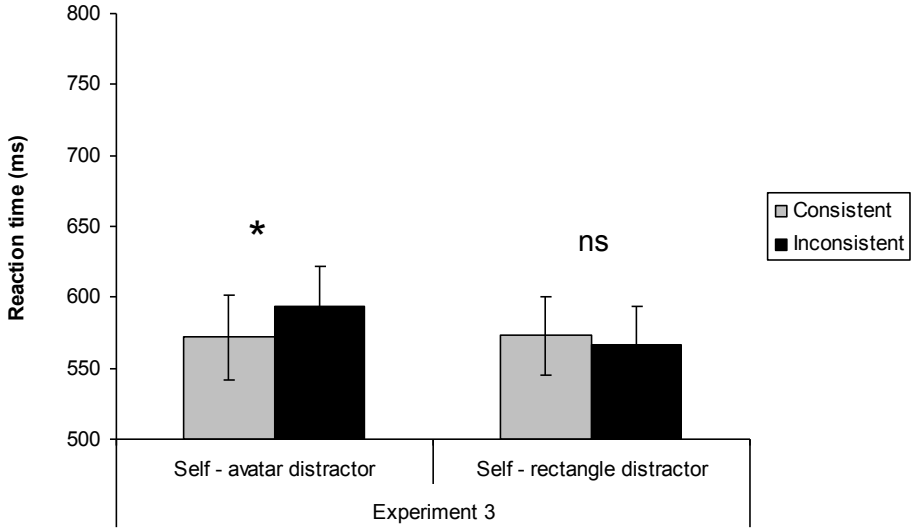
Evidence that mindreading is an efficient but inflexible processes?

- Evidence of involuntary inference of:
  - Simple visual perspective (Samson et al., 2010)

Automatic perspective-taking?

(Samson, Apperly, Braithwaite et al., 2010, *JEP:HPP*)





Only ever judge "self" – how many dots *you* can see

Evidence that mindreading is an efficient but inflexible processes?

- Evidence of involuntary inference of:
  - Simple visual perspective (Samson et al., 2010)
  - Agent's spatial frame of reference (Zwickell, 2011)
  - Agent's "false belief" (Kovacs et al., 2010)

- Sometimes without explicit awareness
  - Schneider et al. (2011)

- Without need for "executive control"
  - Qureshi et al. (2010)

Metacognition: ‘knowledge and cognition about cognitive phenomena’

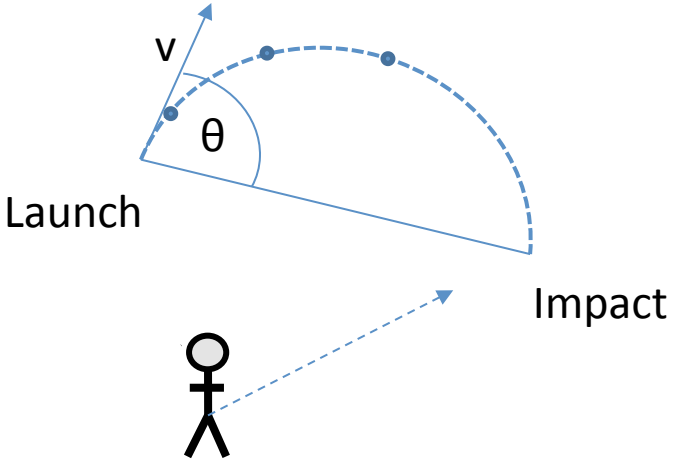
(Flavell 1979, p. 906)

--- e.g. knowledge of others’ beliefs

Physical cognition: knowledge and cognition about physical phenomena

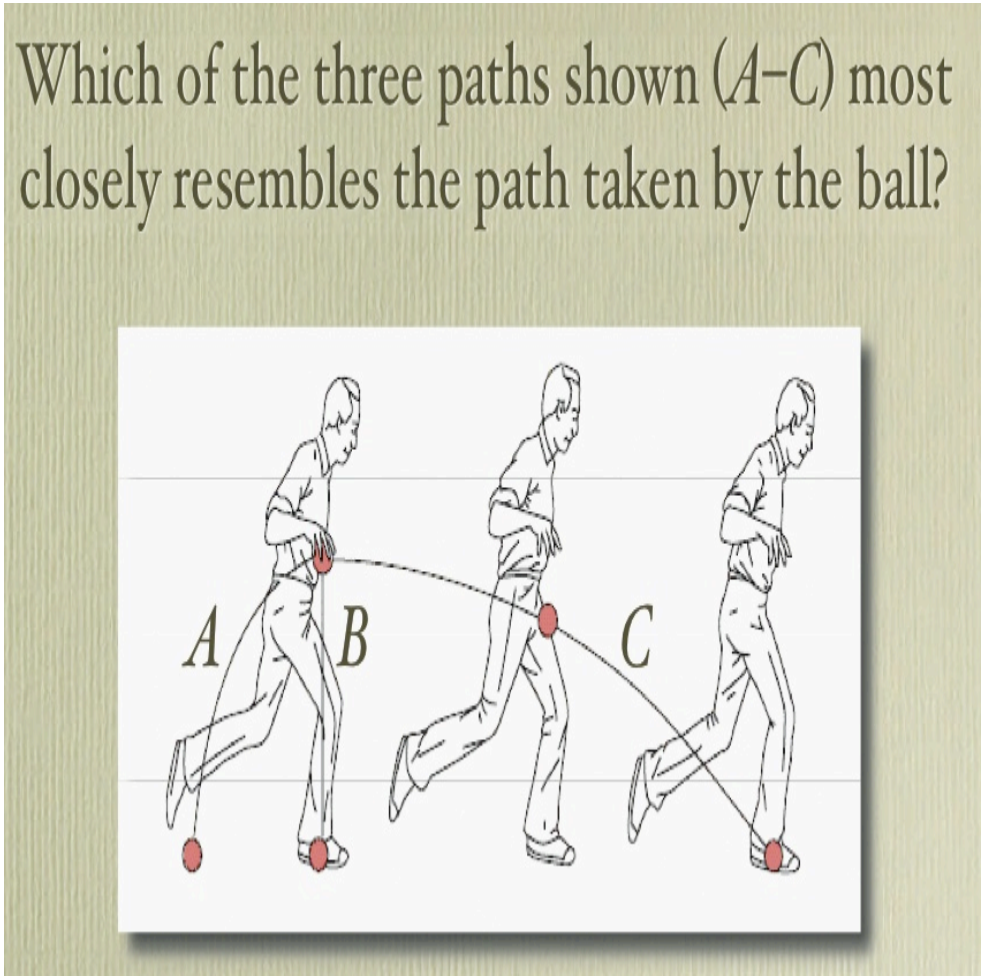
--- e.g. knowledge of trajectories

Examples from the psychology of trajectories



What Newton would have done.....
A) Derive equation for trajectory of ball.
B) Derive equation for one's own capacity to move.
Solve A and B simultaneously

Examples from the psychology of trajectories



McCloskey, Intuitive Physics, Scientific American 248 (1983),

Examples from the psychology of trajectories

- Newtonian mechanics has much greater expressive powers and generalisability than naïve physics

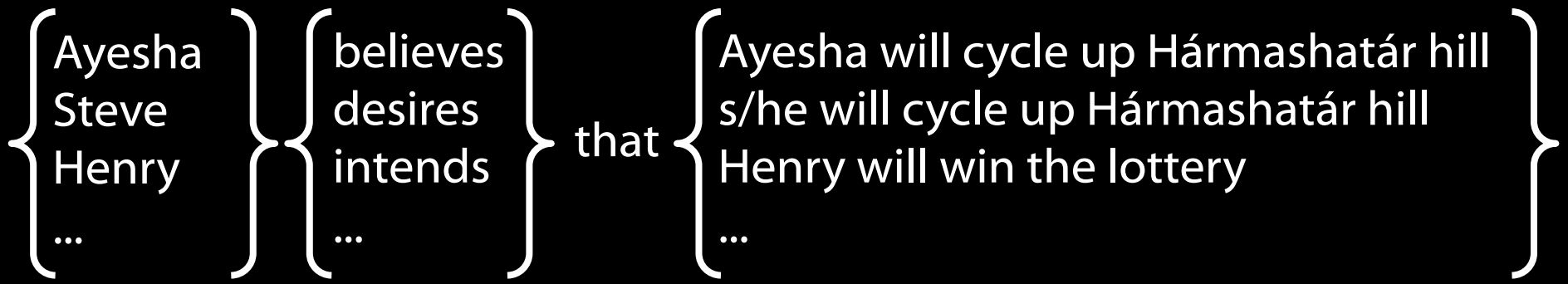
- But this comes at the expense of being slower and more cognitively costly to use

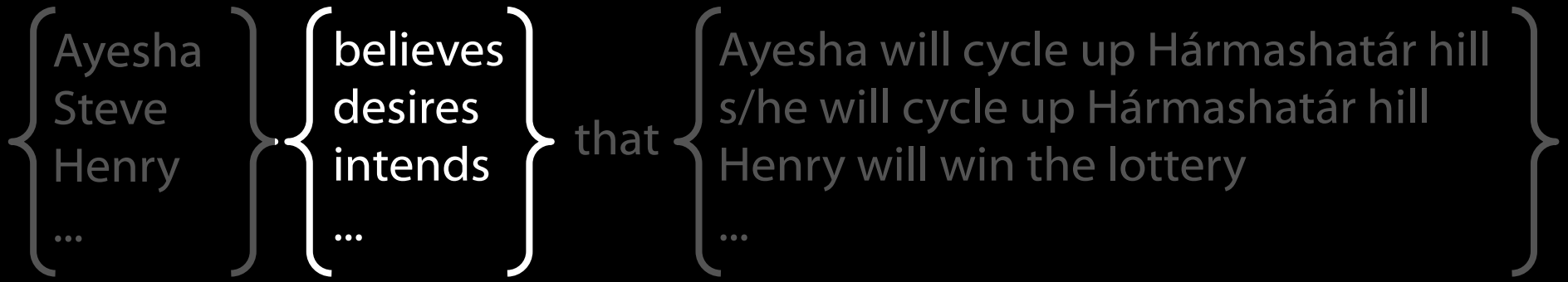
How do we predict his judgment about the falling objects?

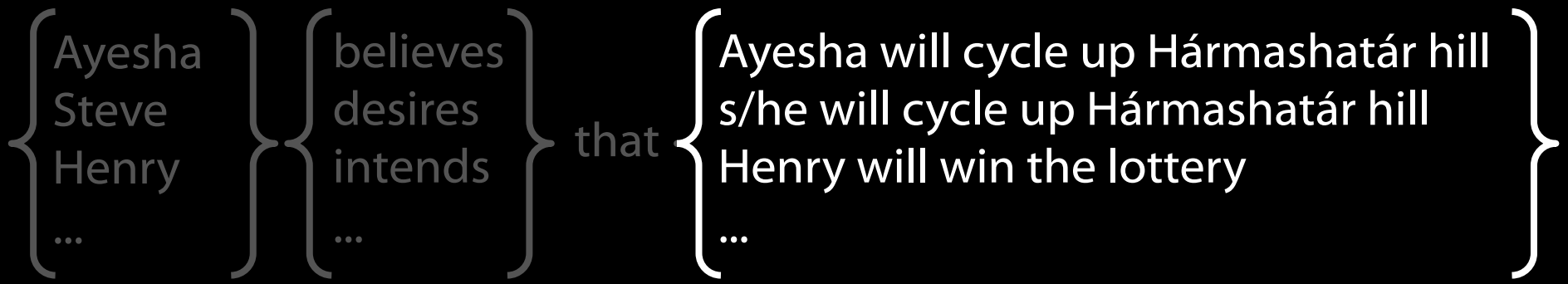
The best theory of trajectories is only of limited use

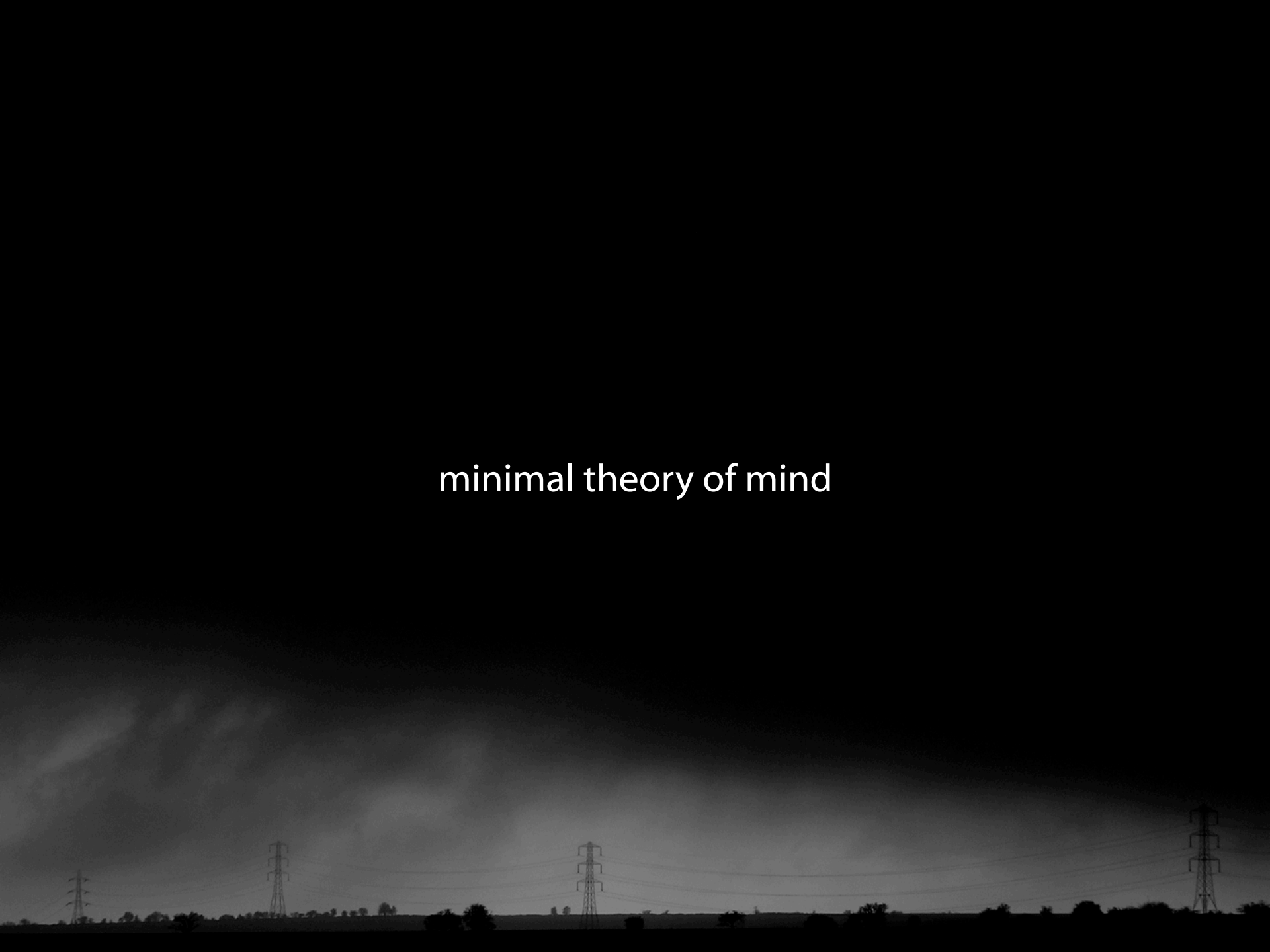
- …for two rather different reasons
  - He may not have that theory
  - Even if he has that theory, it may be too difficult to use in real time

- By analogy, the best theory of mental states may be of limited use for understanding the psychology of mindreading


Ayesha
Steve
Henry
...
believes
desires
intends
...
that
Ayesha will cycle up Hármashatár hill
s/he will cycle up Hármashatár hill
Henry will win the lottery
...
Subject
Attitude
Content

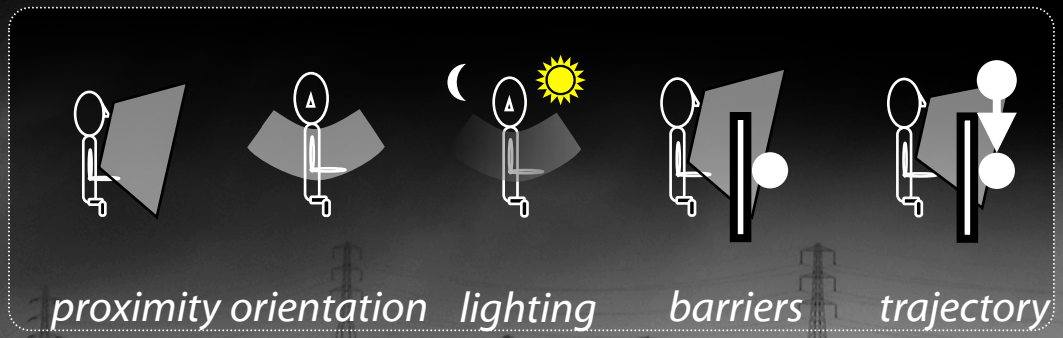

Ayesha
Steve
Henry
...
believes
desires
intends
...
that
Ayesha will cycle up Hármashatár hill
s/he will cycle up Hármashatár hill
Henry will win the lottery
...
Subject
Attitude
Content


Ayesha
Steve
Henry
...
believes
desires
intends
...
that
Ayesha will cycle up Hármashatár hill
s/he will cycle up Hármashatár hill
Henry will win the lottery
...
Subject
Attitude
Content


minimal theory of mind

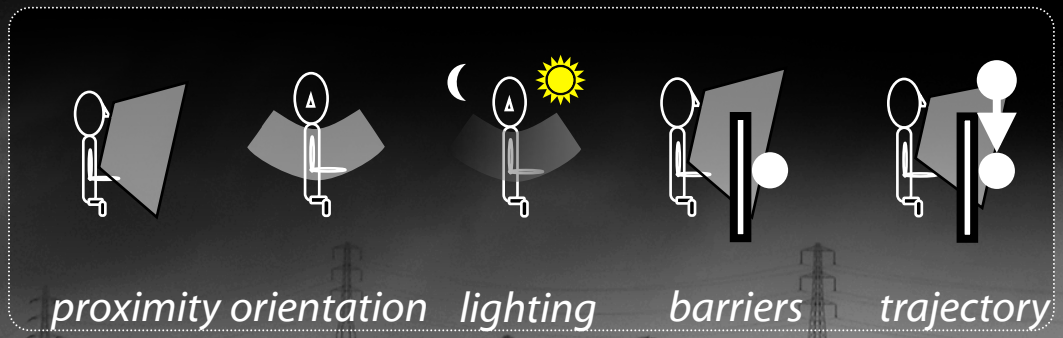
Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

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



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

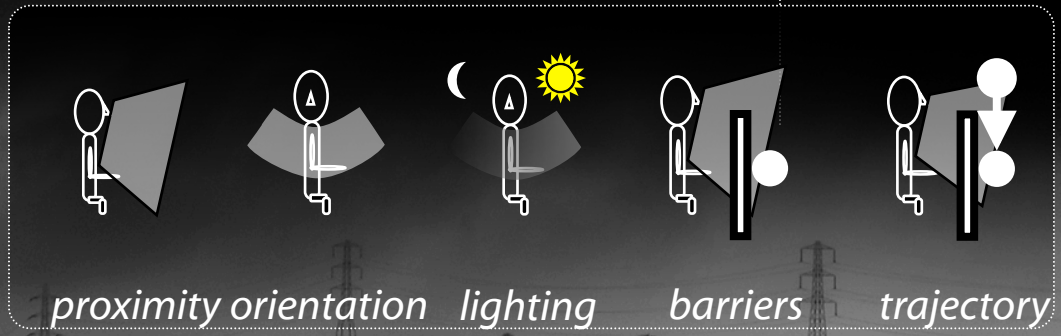
You *encounter* an object = it is in your field

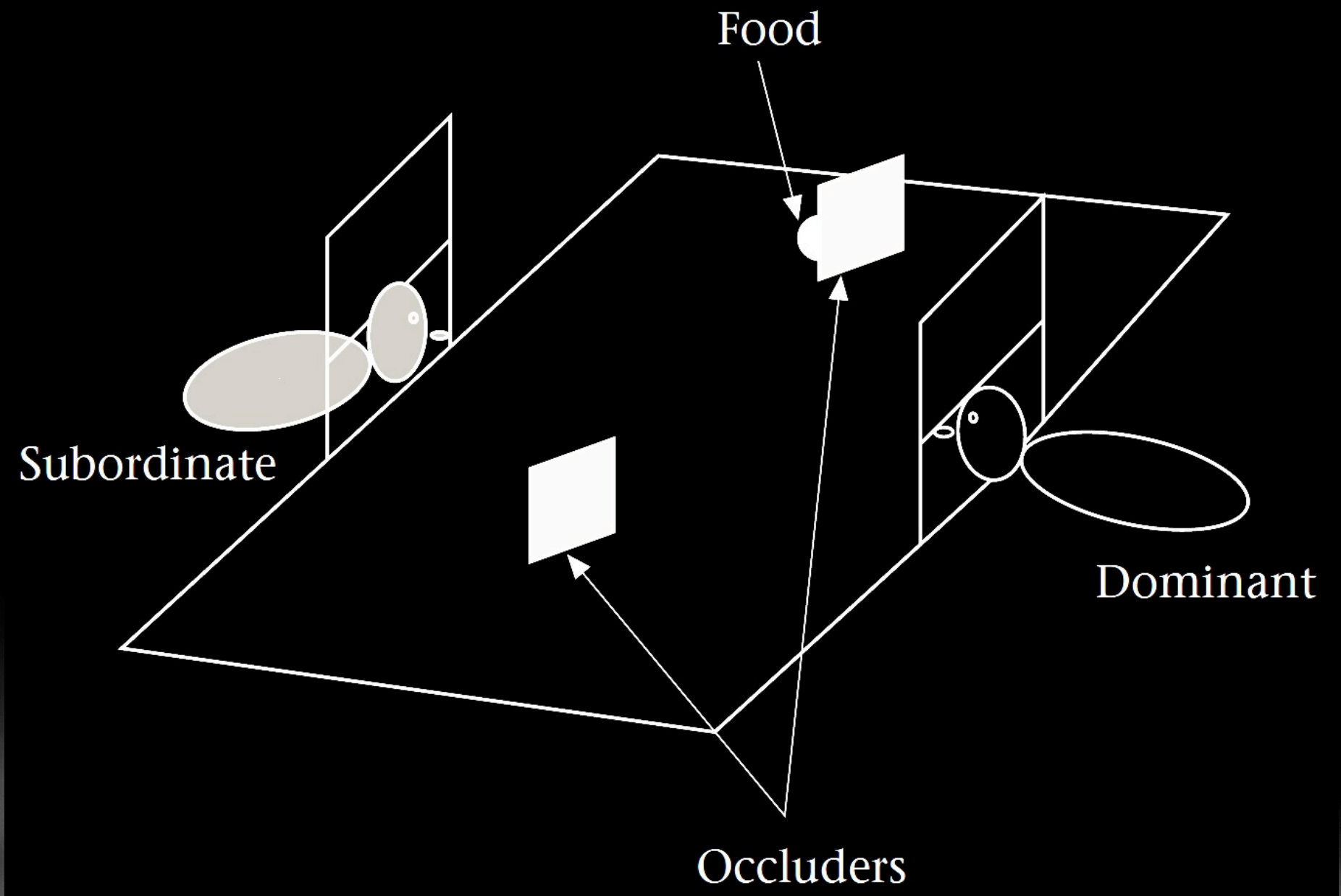


Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

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

Principle 1: one can't goal-directedly act on an object unless one has encountered it.

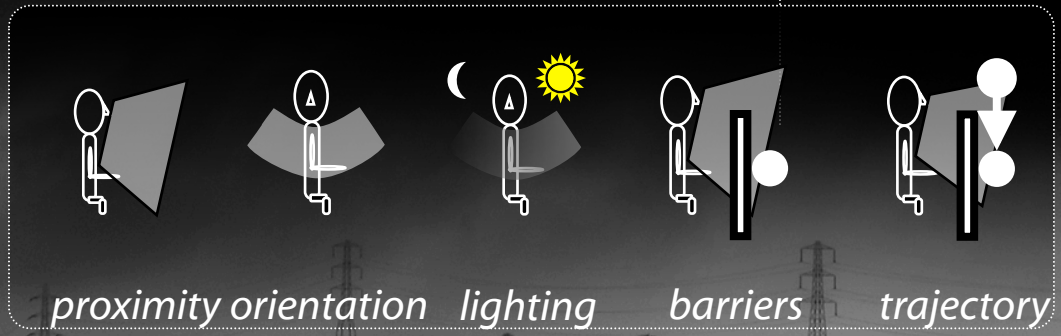



Food
Subordinate
Dominant
Occluders

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

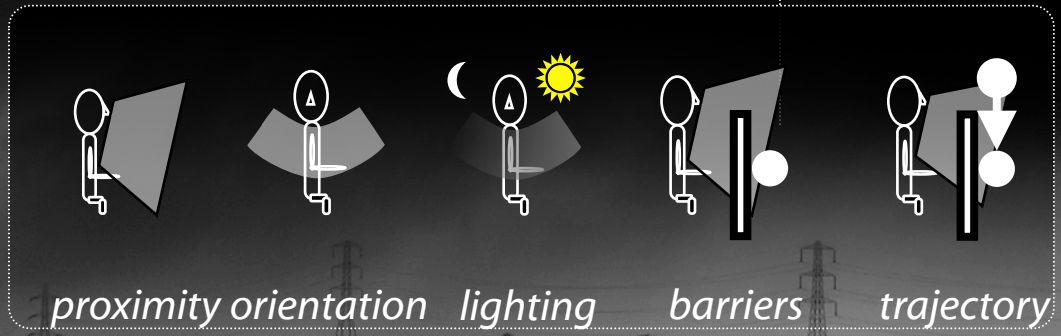


Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

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



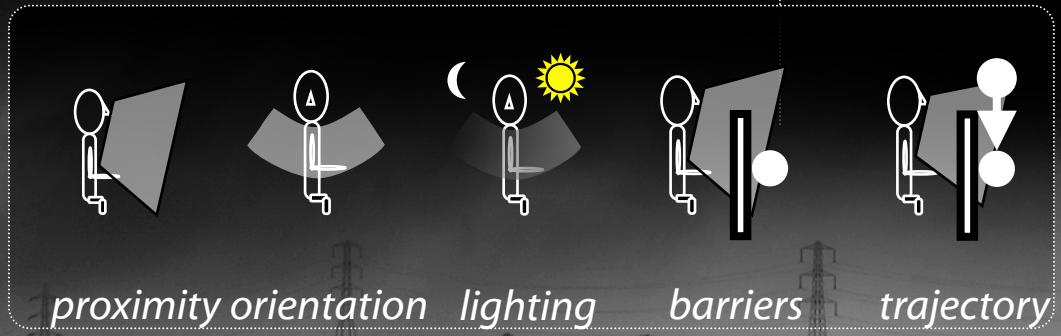
Your *field* = a set of objects related to you by proximity, orientation, lighting and other factors

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

Principle 2: correct registration is a condition of *successful* action.



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

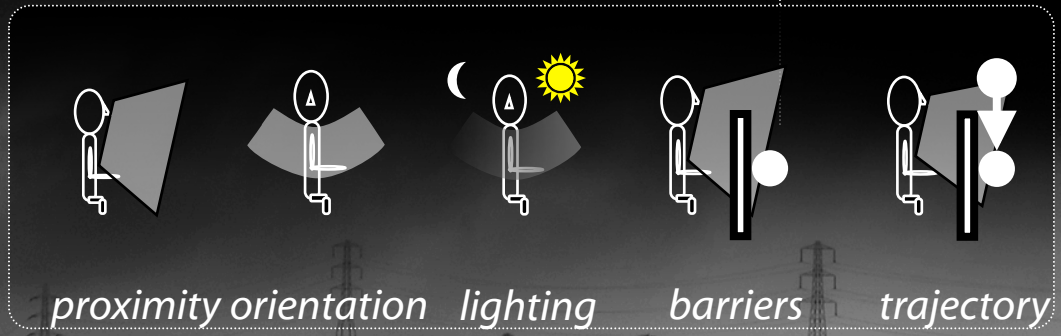
You *encounter* an object = it is in your field

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

Principle 1: one can't goal-directedly act on an object unless one has encountered it.

Principle 2: correct registration is a condition of *successful* action.

Principle 3



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

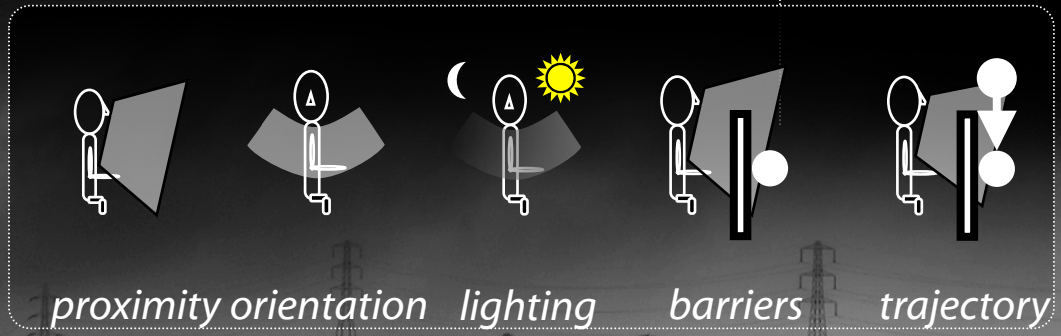
You *encounter* an object = it is in your field

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

Principle 1: one can't goal-directedly act on an object unless one has encountered it.

Principle 2: correct registration is a condition of *successful* action.

Principle 3: 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.



Propositional attitude

Propositional attitude

Relational attitude

Propositional attitude

e.g. believes that ...
e.g. intends that ...
e.g. knows that ...

Relational attitude

e.g. excited by ...
e.g. encountered ...
e.g. wants apple juice

Propositional attitude

e.g. believes that ...
e.g. intends that ...
e.g. knows that ...

arbitrarily nestable contents

uncodifiably complex effects on action

permit mistakes about appearance, identity and existence

Relational attitude

e.g. excited by ...
e.g. encountered ...
e.g. wants apple juice

no contents

parameter-setting effects on action

enable tracking a limited range of false beliefs only

Propositional attitude	Relational attitude
e.g. believes that ... e.g. intends that ... e.g. knows that ...	e.g. excited by ... e.g. encountered ... e.g. wants apple juice
arbitrarily nestable contents	no contents
uncodifiably complex effects on action	parameter-setting effects on action
permit mistakes about appearance, identity and existence	enable tracking a limited range of false beliefs only

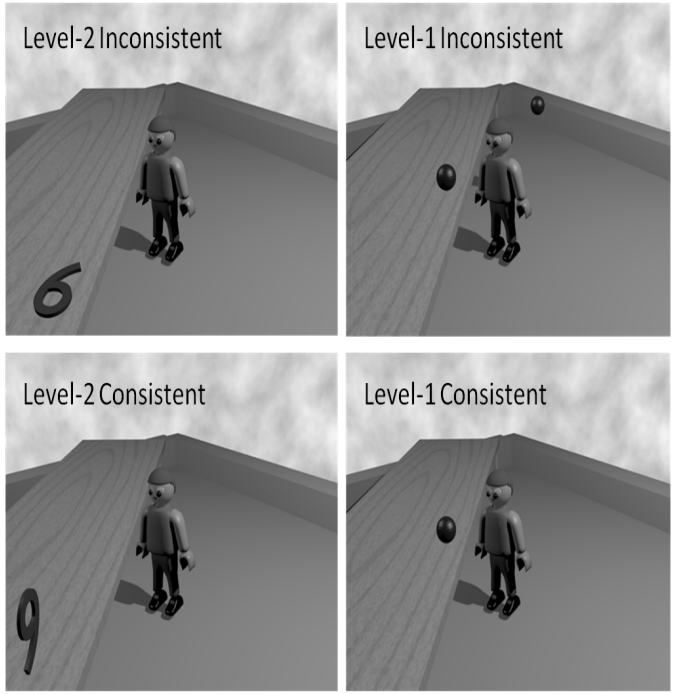
	Propositional attitude	Relational attitude
level-1 perspective taking	Y	Y
level-2 perspective taking	Y	N

	Propositional attitude	Relational attitude
level-1 perspective taking	Y	Y
level-2 perspective taking	Y	N
false beliefs about non-existence	Y	N

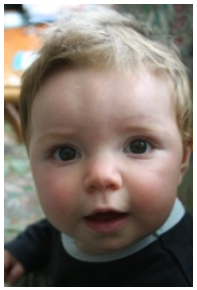
	Propositional attitude	Relational attitude
level-1 perspective taking	Y	Y
level-2 perspective taking	Y	N
false beliefs about non-existence	Y	N
false beliefs about location	Y	Y
false beliefs about identity	Y	N

Evidence that mindreading is an efficient but inflexible processes?

- Evidence of involuntary inference of:
  - Simple visual perspective (Samson et al., 2010)
  - Agent's spatial frame of reference (Zwickell, 2011)
  - Agent's "false belief" (Kovacs et al., 2010)
- Sometimes without explicit awareness
  - Schneider et al. (2011)
- Without need for "executive control"
  - Qureshi et al. (2010)
- Limited to simple cases
  - Level 1 but not Level 2 visual perspectives (Surtees, Butterfill & Apperly, 2012)
  - "False beliefs" about location but not identity (Low & Watts, in press)

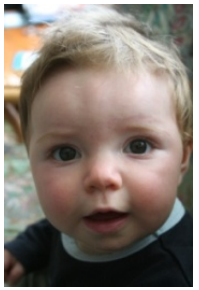


A programme of experimental work



Who is a mindreader?

A programme of experimental work



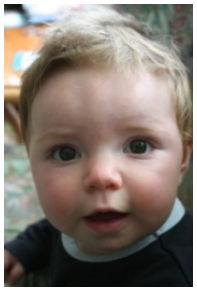
Who is a mindreader?

How does the mindreader model minds?

A programme of experimental work

Present debates assume mindreading requires representing representations, and see FB tasks as a litmus test

There is more than one way of being sensitive to the mental states of others, and it may be very difficult to distinguish between alternatives



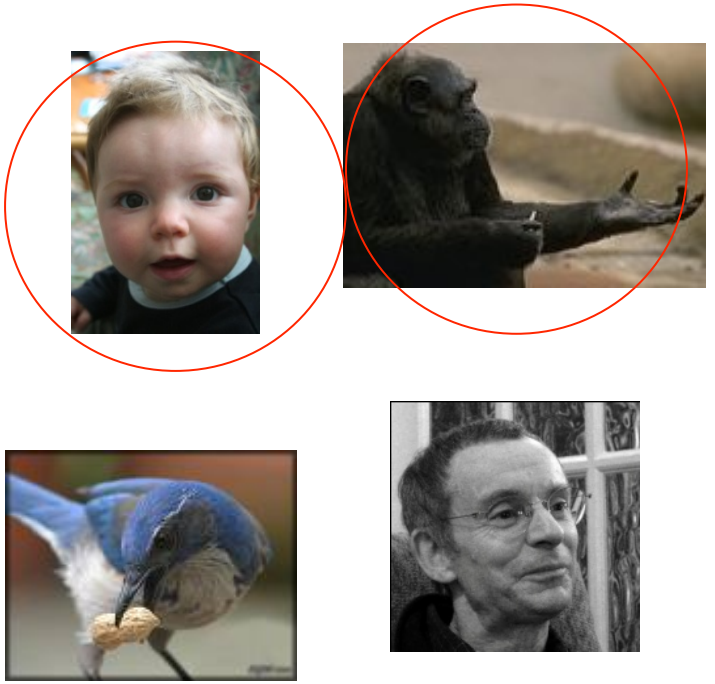
Who is a mindreader?

How does the mindreader model minds?

A programme of experimental work

Present debates assume mindreading requires representing representations, and see FB tasks as a litmus test

There is more than one way of being sensitive to the mental states of others, and it may be very difficult to distinguish between alternatives



Suppose neither could track FB about identity?

Who is a mindreader?

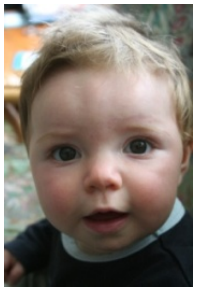
How does the mindreader model minds?

A programme of experimental work

Present debates assume mindreading requires representing representations, and sees FB tasks as a litmus test

There is more than one way of being sensitive to the mental states of others, and it may be very difficult to distinguish between alternatives

- Understanding the *limits* on a given capacity can:
  - Explain how efficiency is achieved
  - Distinguish between alternative mindreading solutions
  - Act as signatures for identifying the operation of a given capacity, across contexts and across types of participant



Who is a mindreader?

How does the mindreader model minds?

Chris' 4th question:

What advantages are conferred by explicit mindreading?

- Explicit mindreading escapes hard limits
- It does so at the expense of being cognitively costly
- (But if we believe explicit mindreading entails abductive inference, simply noting that it is demanding of resources for memory and cognitive control does not actually explain how it is possihble)

engineers not scientists

1. Is there implicit, i.e. unconscious, meta-cognition?

1. Is there implicit, i.e. unconscious, meta-cognition?

3. Is there a systematic difference between the content of implicit and explicit meta-cognition?

1. Is there implicit, i.e. unconscious, meta-cognition?

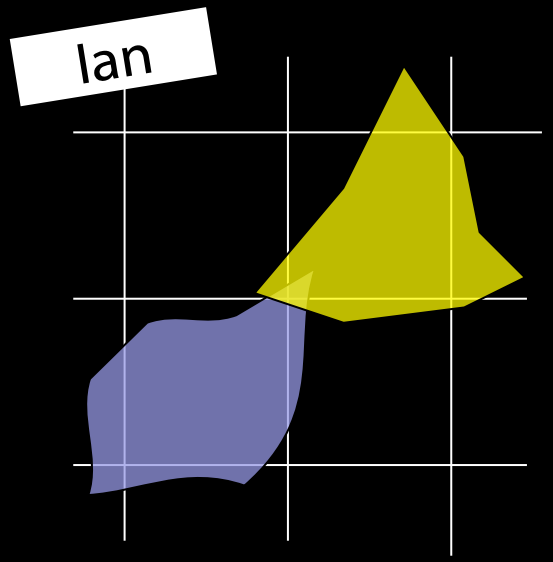
2. When we track others' mental states without awareness of doing so, should these implicit processes be considered meta-cognitive?

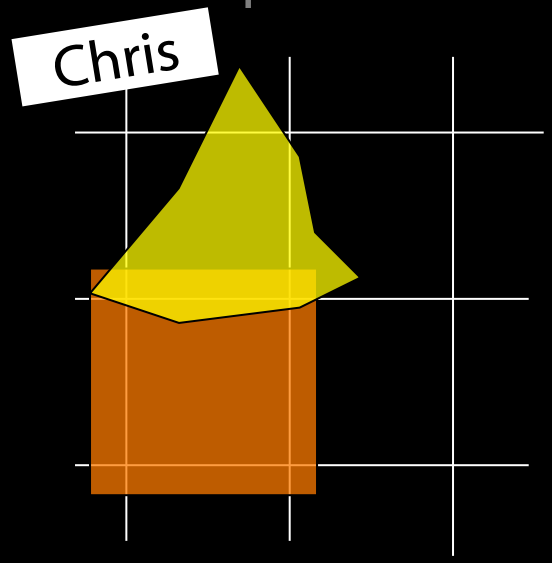
3. Is there a systematic difference between the content of implicit and explicit meta-cognition?

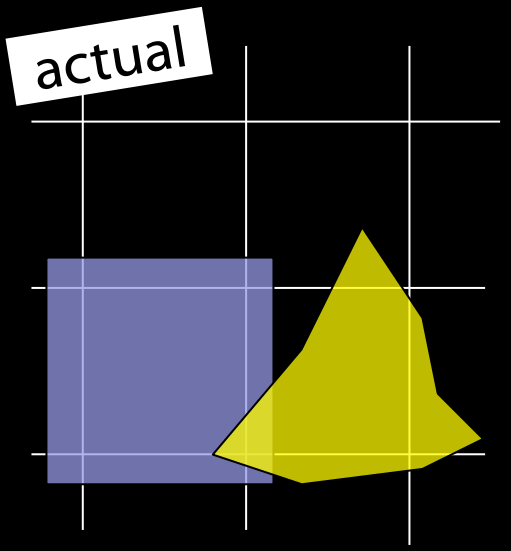
1. Is there implicit, i.e. unconscious, meta-cognition?

2. When we track others' mental states without awareness of doing so, should these implicit processes be considered meta-cognitive?

3. Is there a systematic difference between the content of implicit and explicit meta-cognition?







Metacognition: 'knowledge and cognition about cognitive phenomena'

(Flavell 1979, p. 906)

2. When we track others' mental states without awareness of doing so, should these implicit processes be considered meta-cognitive?

3. Is there a systematic difference between the content of implicit and explicit meta-cognition?