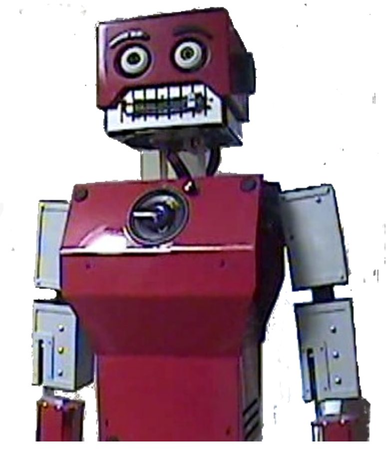
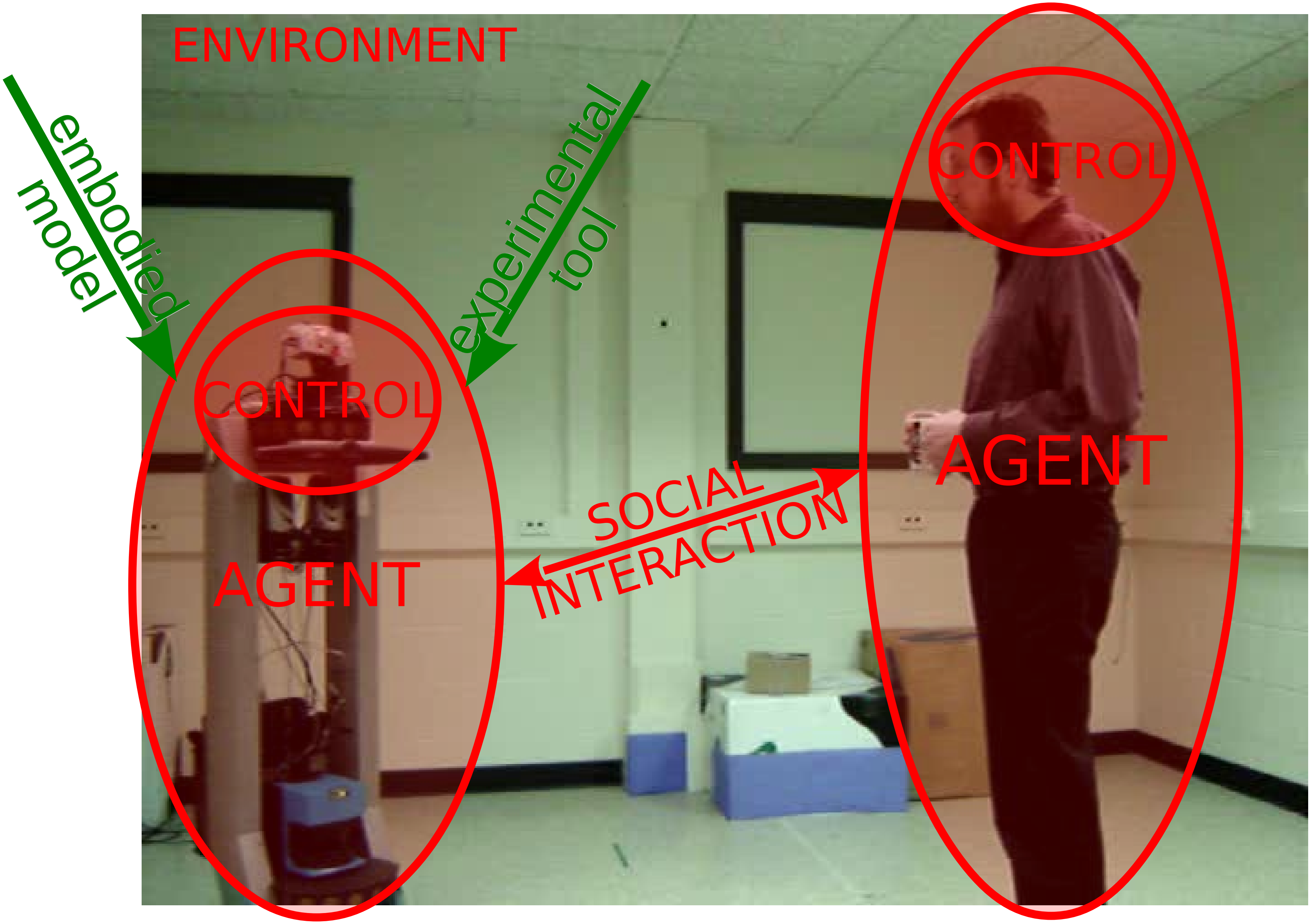
Human-Robot Interaction:
Where Cognitive Science Meets Computer Science and Robotics

Matthias Scheutz
matthias.scheutz@tufts.edu
Human-Robot Interaction Laboratory
Department of Computer Science
Tufts University
Medford, MA 02155
http://hrilab.tufts.edu/



Human-robot interactions



The two main uses of robots in cognitive science

Robots have a dual role in cognitive science as

- **tools for studying situated human cognition**
  - *non-cognitive robots*

  - *cognitive robots*

Examples of robots as tools

- Joint attention processes
  (e.g., establishing and maintaining joint attention, or it breaking joint attention through “abnormal attention”)
- Human attitudes about robots
  (e.g., social facilitation and social inhibition to probe agency, or investigations of the effects of robotic voices, social presence, etc.)
- Human reactions to autonomous robots
  (e.g., to robot affect, robot autonomy, to local/remote HRI)
- Task-switching in human multi-tasking
  (e.g., fNIRs-based adaption of robot autonomy, effects of real vs virtual robots on multi-tasking performance)
- Philosophical and conceptual inquiry
  (e.g., what it is like to be an agent/have a color experience, or the effects of “ethical robots” on human decision-making)

The two main uses of robots in cognitive science

Robots have a dual role in cognitive science as

- **tools for studying situated human cognition**
  - *non-cognitive robots*
  - *cognitive robots*
- **embodied models of cognition**
  - *non-interaction models*
  - *interaction models*

Examples of robots as models

- Spoken natural language and dialogue interactions
  (e.g., instructing and tasking in natural language, dialogue-based mixed initiative, robust NL interactions under time pressure)
- Introspection and self-awareness
  (e.g., detecting faults and failures, detecting capabilities, automatic adaptation of architectural components for improved autonomy)
- Planning, reasoning, and problem solving in “open worlds”
  (e.g., planning and reasoning with incomplete knowledge, determining optimal policies in open worlds)
- Knowledge-based interactive learning
  (e.g., one-shot learning of new actions, new plan operators, and new perceivable objects)
- Mental models, simulation, and counterfactual reasoning
  (e.g., adverbial cues for inferring false beliefs, automatic inference from mental models, simulations of actions)

What do we need?

- To employ robots in both roles, we need the right kind of **computational framework** in which we can develop both interaction experiments and computational models
- Over the last decade, we have developed such a framework which consists of two parts:
  - **DIARC** – a "Distributed Integrated Reflective Affective Deliberative" architecture framework (e.g., Scheutz et al. 2013, Cantrell et al. 2010, Scheutz et al. 2010, Schermerhorn and Scheutz 2010, Scheutz et al., 2007 Schermerhorn et al. 2005)
  - **ADE** – the "Agent Development Environment" middleware (e.g., Scheutz 2006, Kramer and Scheutz 2007, and others)
- DIARC is implemented in ADE and consists of several specific architectural control components that implement different cognitive functions (some of which are biologically plausible, while others are engineering solutions to enable and/or facilitate the development of integrated models)

The two main uses of robots in cognitive science

Robots have a dual role in cognitive science as

- **tools for studying situated human cognition**
  - *non-cognitive robots*
  - *cognitive robots*
- **embodied models of cognition**
  - *non-interaction models*
  - *interaction models*

Four examples

Robots have a dual role in cognitive science as

- **tools for studying situated human cognition**
  - *non-cognitive robots* (e.g., robots acting in an environment to study human attention shifts)
  - *cognitive robots* (e.g., the roles of affect and embodiment in task-based team interactions)
- **embodied models of cognition**
  - *non-interaction models* (e.g., models of conjunctive visual search guided by spoken instructions)
  - *interaction models* (e.g., models of indirect speech acts in dialogue-based interactions)

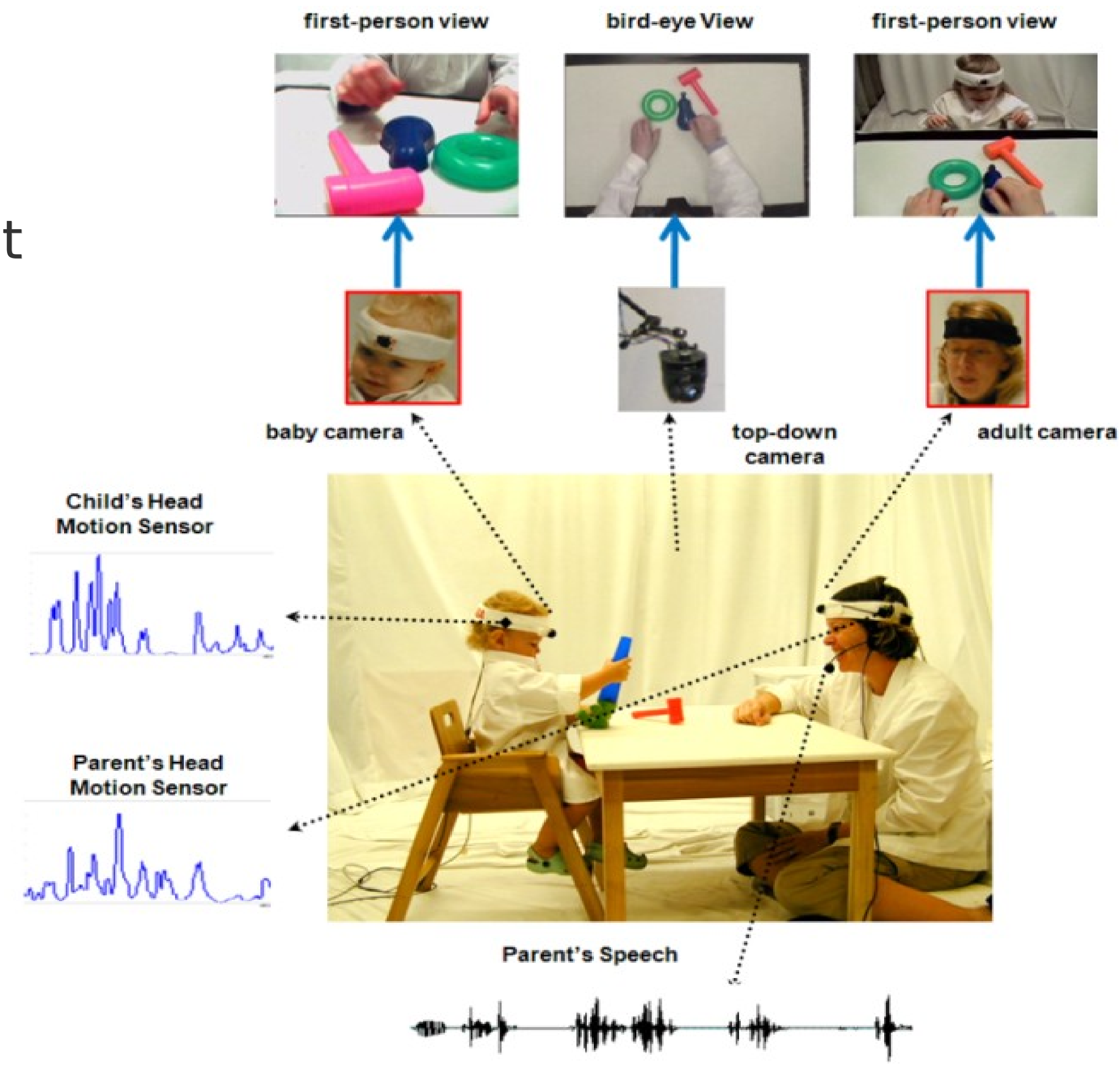
Example 1:
joint attention

Robots have a dual role in cognitive science as

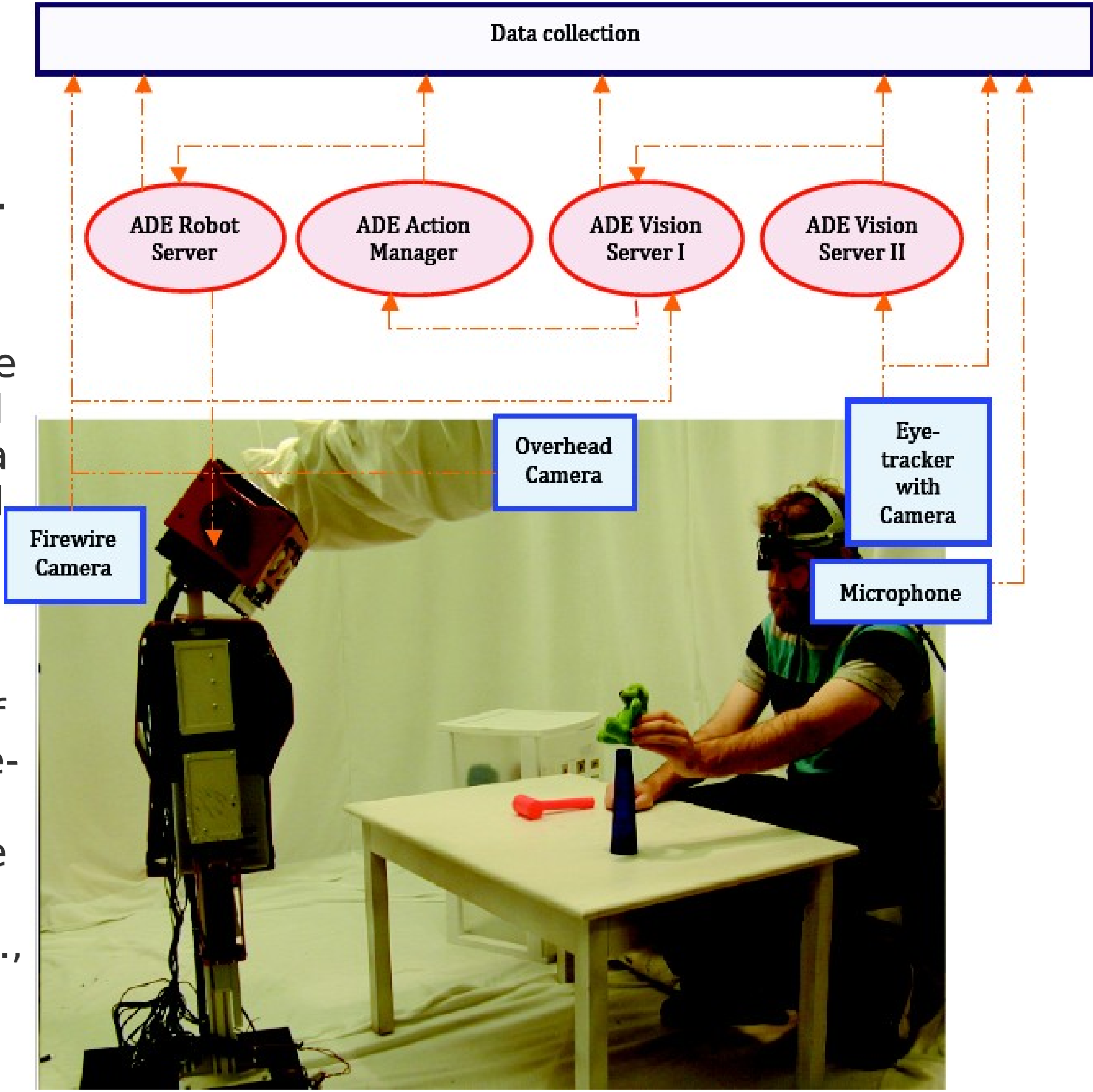
- **tools for studying situated human cognition**
  - *non-cognitive robots* (e.g., robots acting in an environment to study human attention shifts)
  - *cognitive robots* (e.g., the roles of affect and embodiment in task-based team interactions)
- **embodied models of cognition**
  - *non-interaction models* (e.g., models of conjunctive visual search guided by spoken instructions)
  - *interaction models* (e.g., models of indirect speech acts in dialogue-based interactions)

Joint attention processes in parents and children

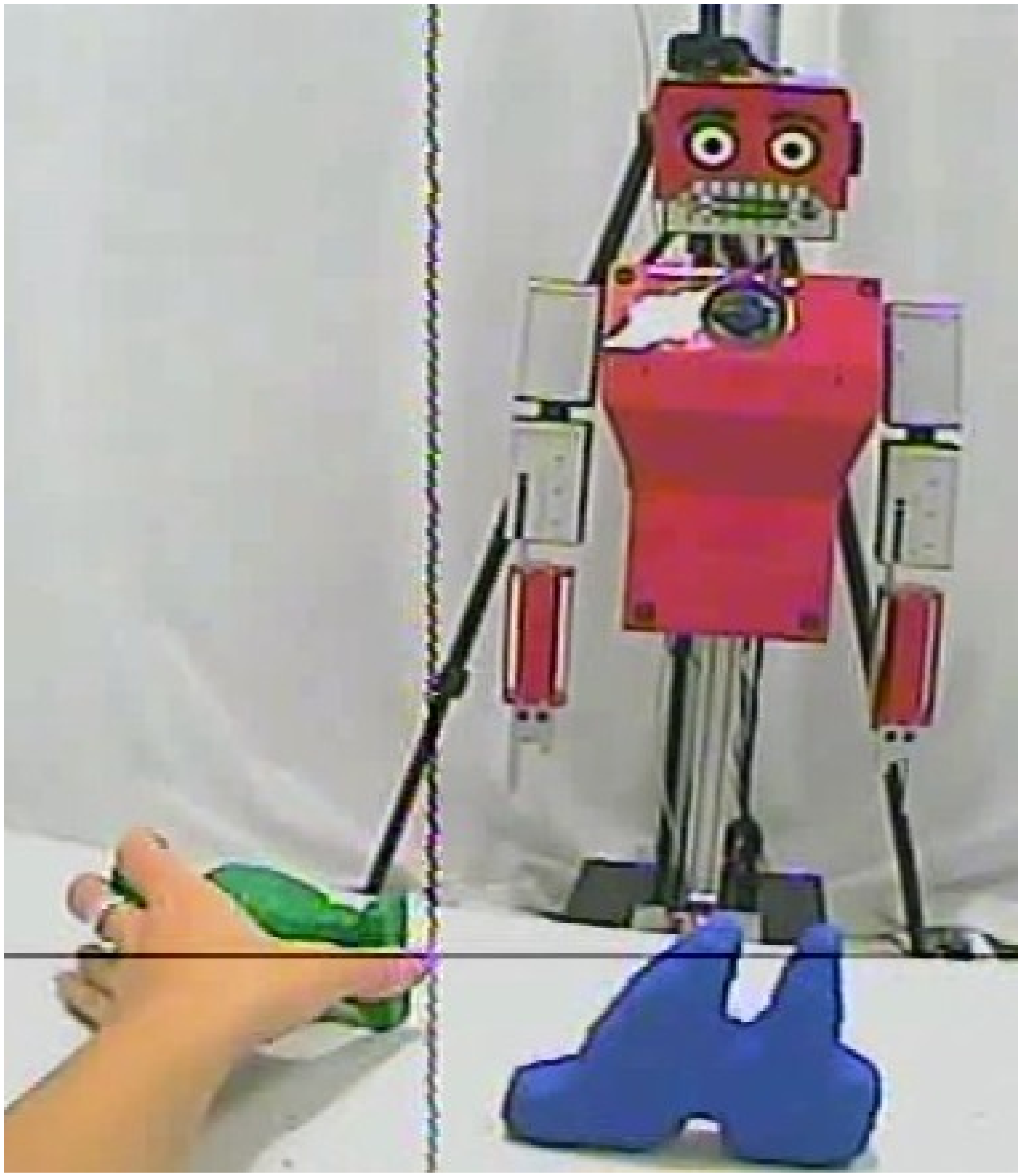
- Yu, Smith, Shen, Pereira, and Thomas (2009) studied the different dynamic structures of children's and parents' views of the events in the shared task of toy play and word learning
- Multi-modal data recording to obtain detailed time-course information



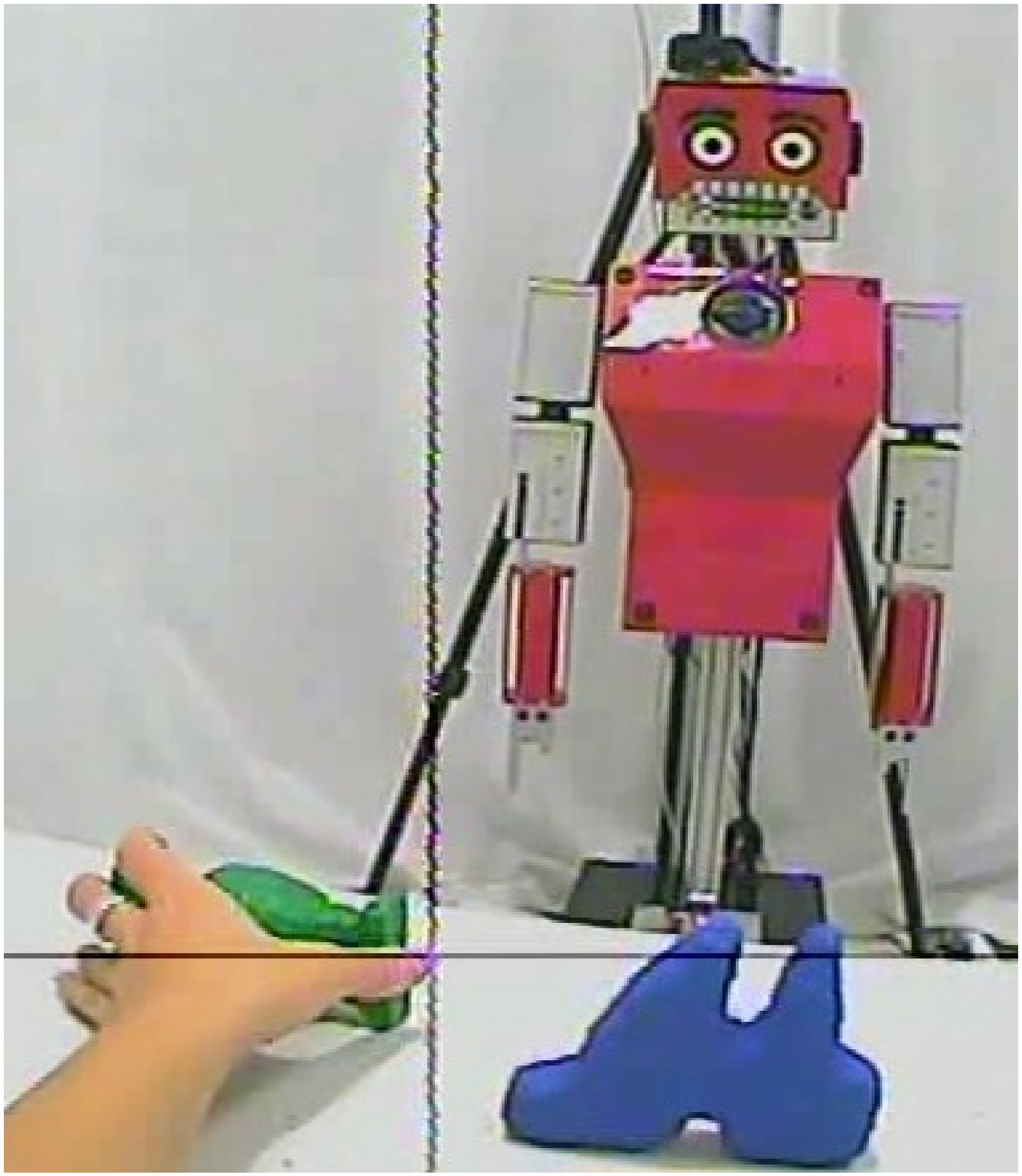
- Yu et al. (2010, 2011) replicated the experimental setup using a robot instead of human participant
- Required processing of real-time eye-gaze data and real-time reaction to the data (e.g., head moves)

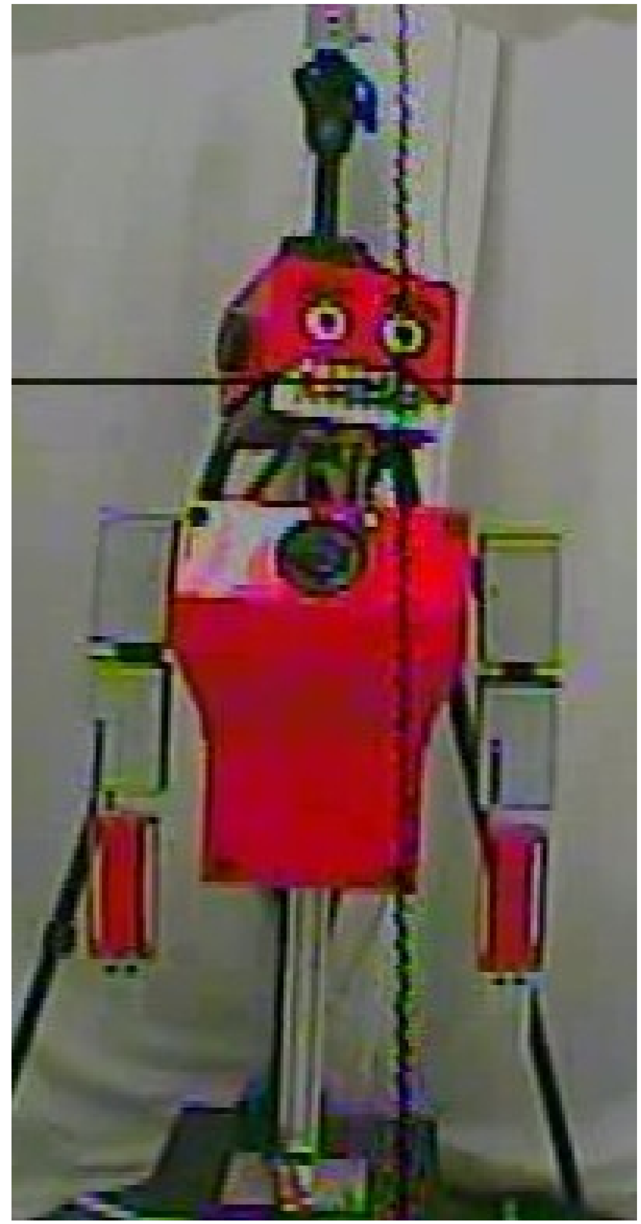


The “follow” condition

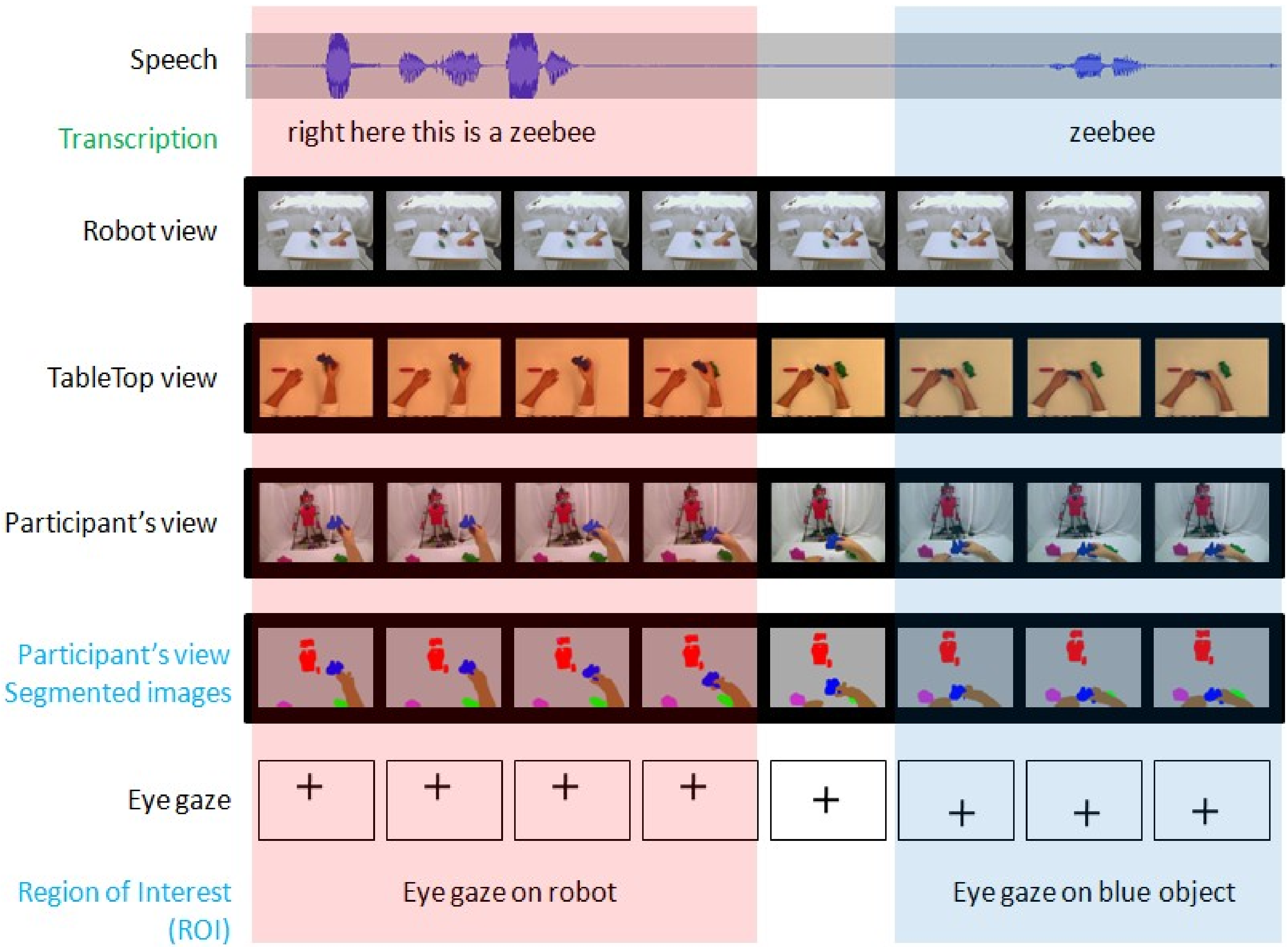


The “random” condition



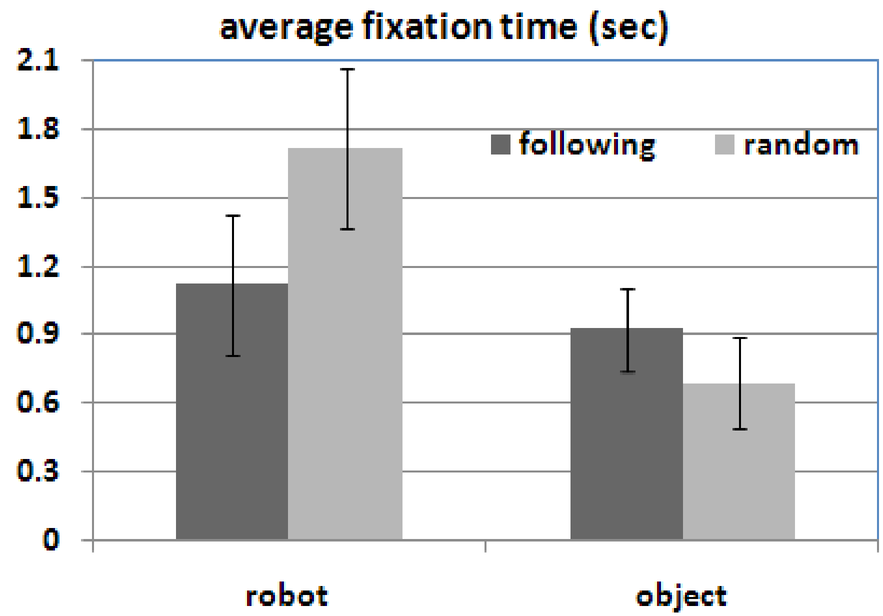


Multi-modal data collection



Eye fixations results

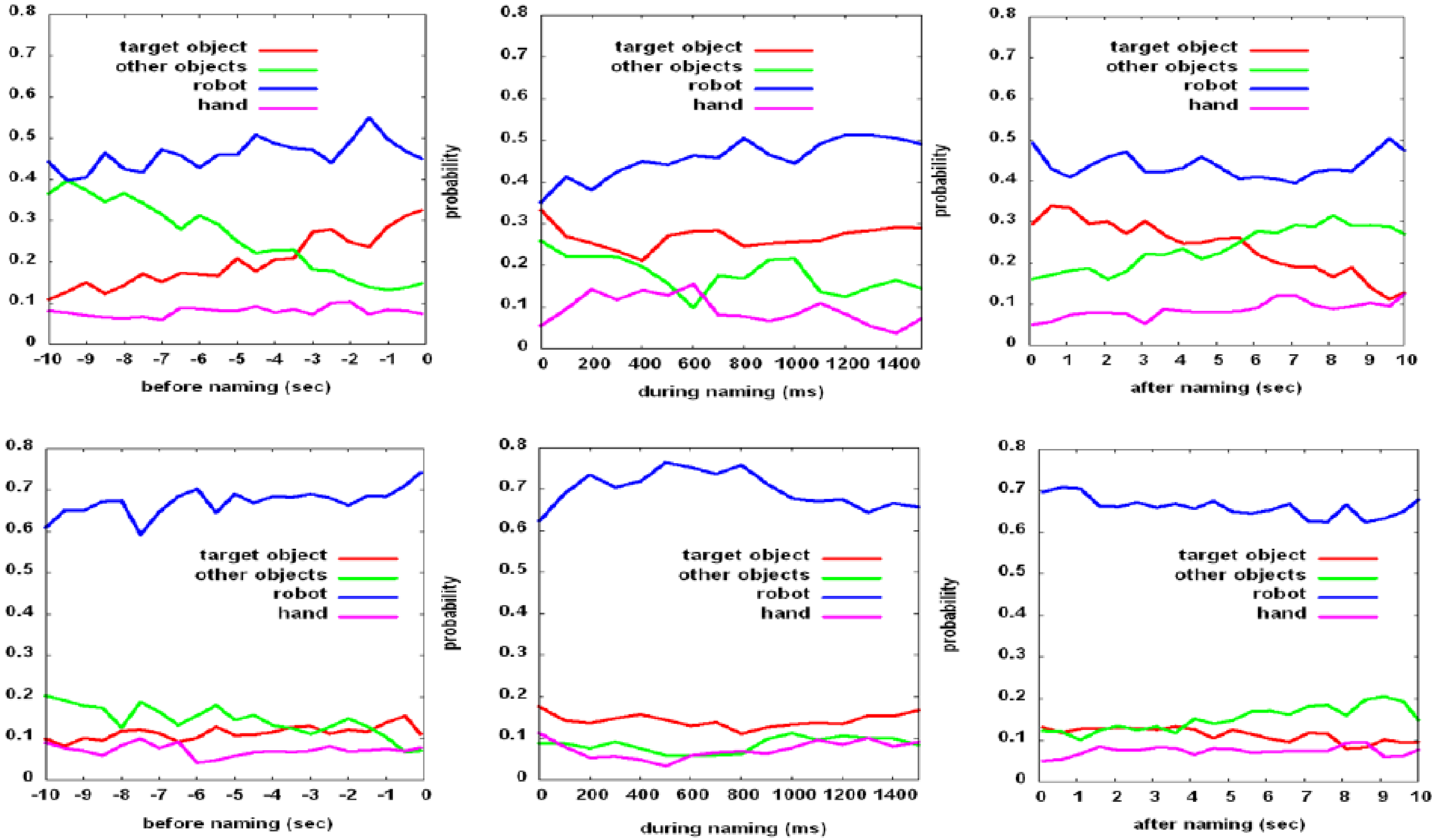
- Participants in the random condition visually attended to the robot significantly longer (through longer eye fixations) than to objects and also longer than those in the following group



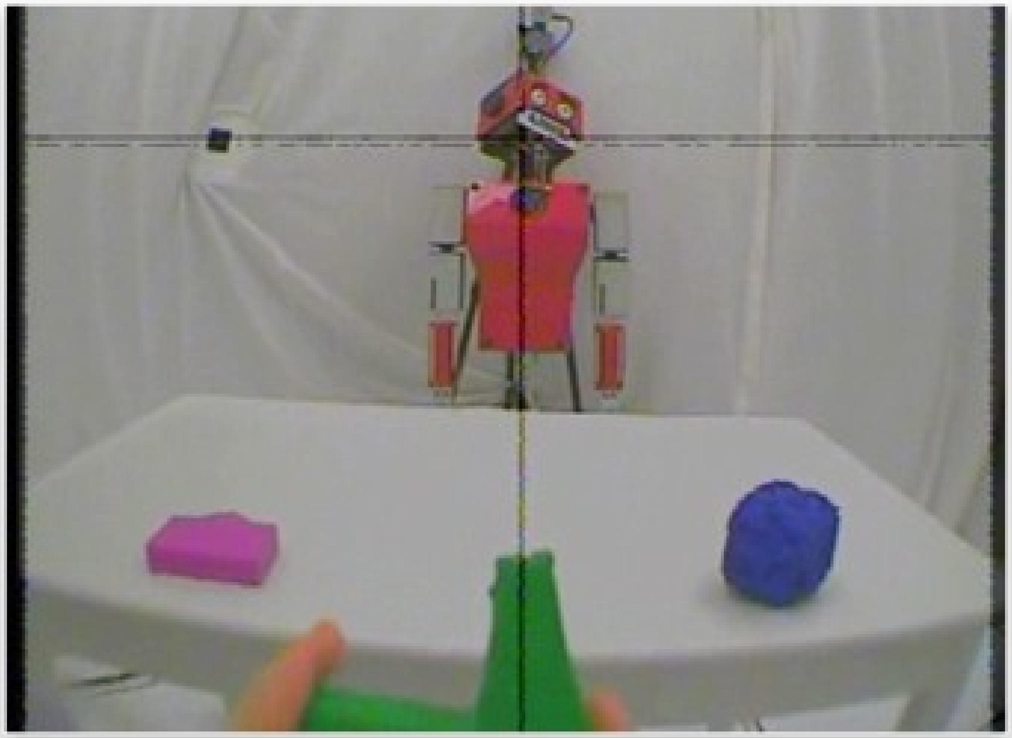
- Participants in the random condition generated more attention-attracting utterances and more naming utterances than participants in the following condition

	following	random
"look"	2.12	6.33
"see"	2.23	7.78
"here"	2.92	11.12
"robot"	0.33	3.44
"hey"	0	1.78
"yes"	1.10	2.83
object names	48.78	60.44

Temporal dynamics before, during and after naming events

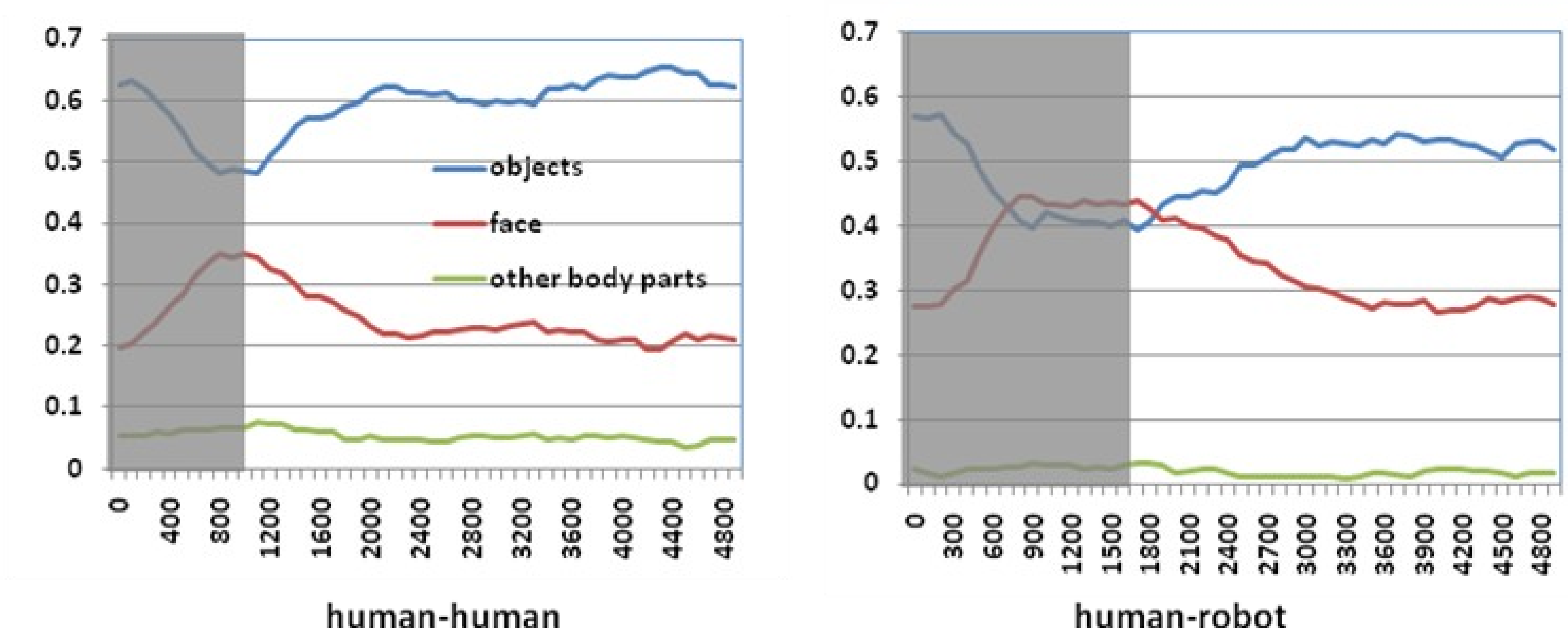


The “human robot” condition



- Use “human robot” (performed by trained actor) to behave exactly like the robot in the “random condition” to be able to better compare the subjects' attention processes across conditions and to study potential differences in appearance
- Use pre-determined sequence of minimal behavioral cues (only head motion to pre-determined positions)

Eye gaze during and after agent head turns



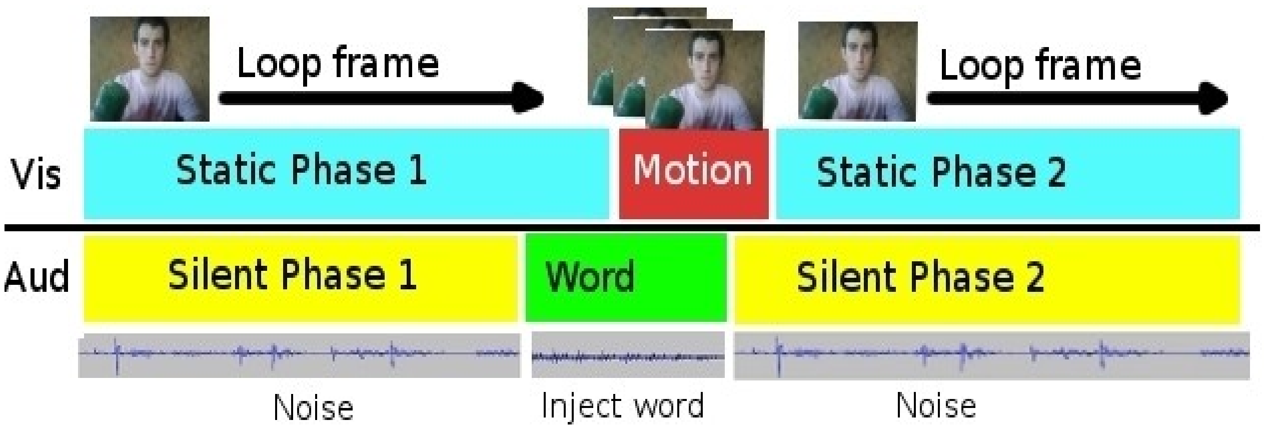
- Note that the robot took a longer time than the human to generate the same head movement
- Nevertheless, the results showed that participants in both conditions quickly switched their attention to the agent's face soon after the onset of the head turn, and then back to the target object right after the offset of the head turn

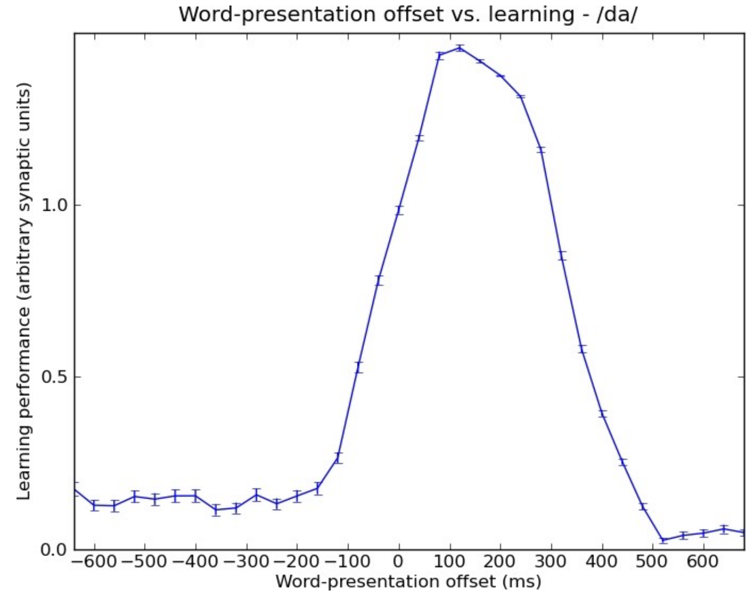
Implications

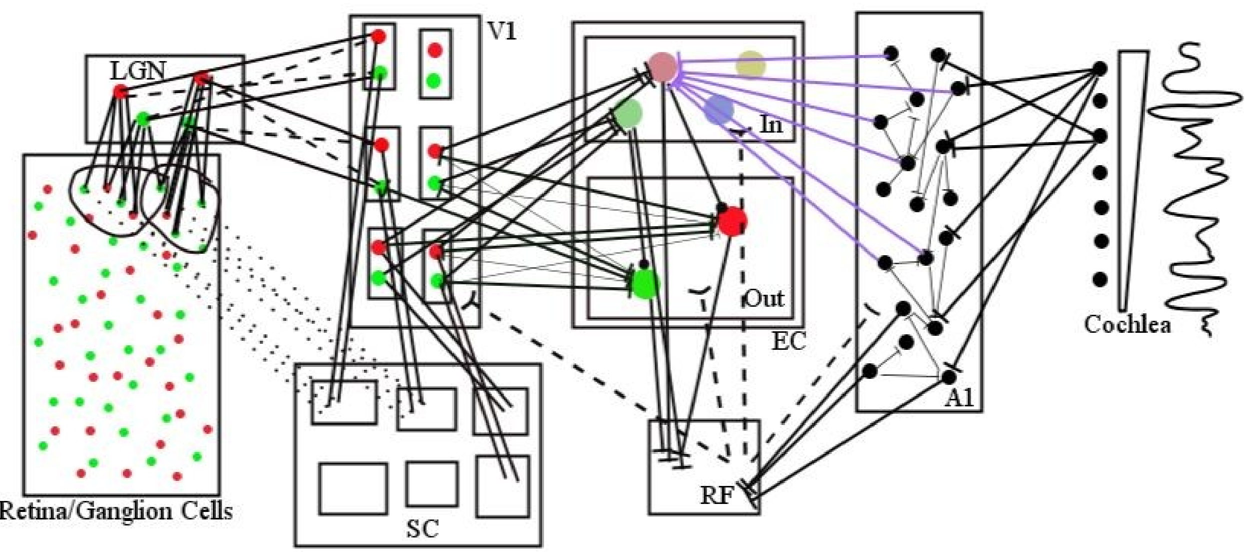
- Human attention allocation is significantly impacted by a robot's joint attention behavior
- While the robot's particular appearance (robotic vs human-like) can have some modulatory effect w.r.t. the overall probablility of eye fixation, the overall effect on human attention allocation is the same
- Also studied gender effects in this task (using a female actor) – won't be able to discuss them here
- Main lesson:

  DON'T PUT EYES ON IT IF IT CAN'T DO WHAT WE EXPECT EYES TO DO!

Joint attention and infant word-object learning







Example 2: affect and autonomy

Robots have a dual role in cognitive science as

- **tools for studying situated human cognition**
  - *non-cognitive robots* (e.g., robots acting in an environment to study human attention shifts)
  - *cognitive robots* (e.g., the roles of affect and embodiment in task-based team interactions)
- **embodied models of cognition**
  - *non-interaction models* (e.g., models of conjunctive visual search guided by spoken instructions)
  - *interaction models* (e.g., models of indirect speech acts in dialogue-based interactions)

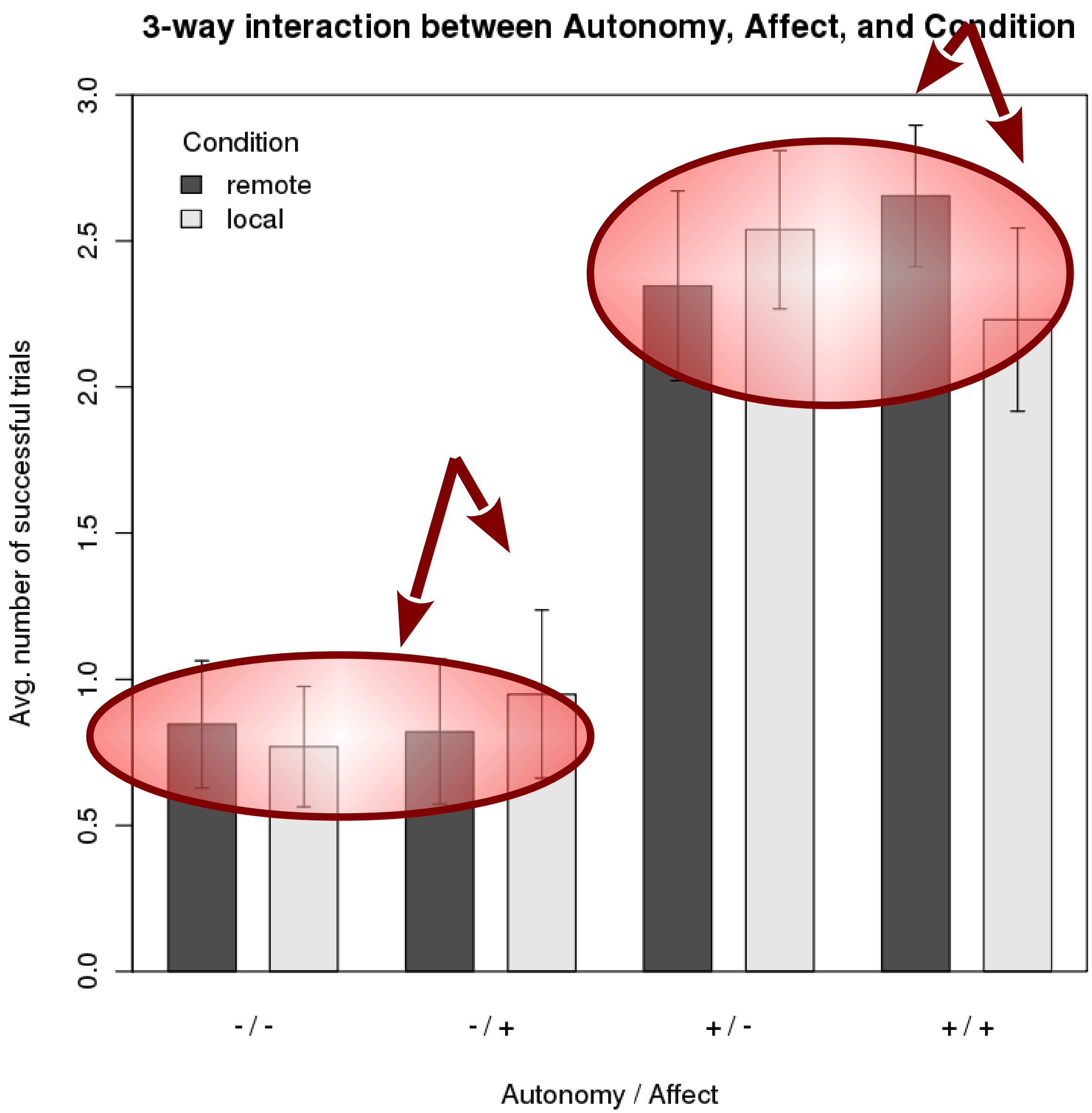
Affect and Embodiment in HRI

- Various studies in HRI have looked at modulatory aspects in human-robot interactions, including the differences between simulated vs. physically co-located robots:
  - a co-located robot was treated in a much more human-like manner compared to a simulated robot (Bainbridge et al. 2008)
  - human interviewees were more engaged and guarded in their disclosures with a physical compared to simulated robot (Kiesler et al. 2006)
- Similarly, a robot's affect expression at key points in a human-robot team task has been shown to led to better performance of the team compared to no affect in the voice (Scheutz et al. 2006, Schermerhorn and Scheutz, 2009, 2011)
- Want to systematically study the tradeoffs between affect and robot embodiment in a collaborative HRI team task

The HRI Exploration Task

Some results (objective measures)

- Overall, **autonomy** subjects were more likely to **successfully complete the transmission** than **no-autonomy** subjects (mean number of successful runs: 2.435 vs. 0.845, p<.001)
- There was a 3-way interaction on **successful transmissions** (p=.0368)
    - post-hoc analysis (Tukey's HSD) confirms that all **autonomy** groups are significantly higher than all **no-autonomy** groups



Some results (subjective measures)

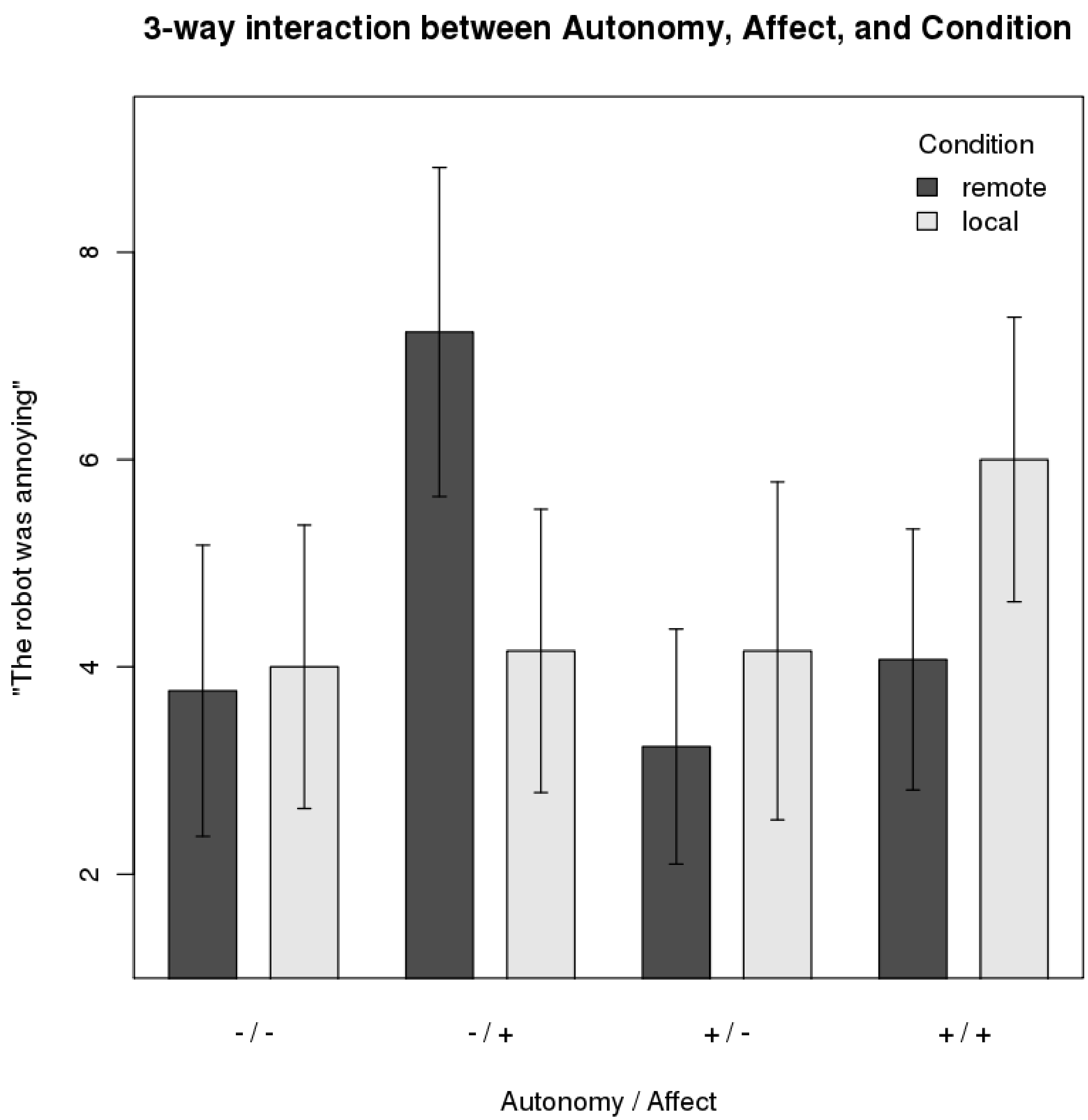
- Subjects in the **affect** group were more likely to agree that
  - the robot “sounded like someone expressing a mood or emotion” (6.376 vs. 3.290, $p<.001$); and that
  - the robot “seemed to have emotions of its own when it spoke” (6.284 vs. 3.288, $p<.001$)
- And even though subjects in the **affect** group rated the robot higher for **responsiveness to commands** (6.311 vs. 5.920, $p=0.0347$) and for **cooperativeness** (5.882 vs. 6.368, $p=.0200$), affect did not influence attitudes about robots in general
- Instead **autonomy** subjects were more likely to agree that
  - “some robots should be capable of behaving convincingly as though they have emotions” (6.170 vs. 5.096, $p=.0140$)
  - “some robots should be capable of making their own decisions” (5.412 vs. 4.177, $p=0.0163$)
  - and even that “some robots should be able to choose to disobey humans in some situations” (5.360 vs. 4.133, $p=.0170$

Some results (subjective measures)

- This is true despite the fact that **autonomy** subjects were more likely to report that the robot "appeared to disobey commands" (4.939 vs. 3.620, $p=.0157$), and that, in fact, the robot *did* disobey in the **autonomy** condition
- Why? "Substance seems to win out over form", for the **autonomy** robots get higher ratings on:
  - acting like a **member of the team** (6.697 vs. 5.822, $p=.0398$)
  - being **easy to interact with** (6.673 vs. 5.654, $p=.0235$)
  - being **helpful** (7.176 vs. 5.305, $p<.001$)
  - being **capable** (7.618 vs. 5.863, $p<.001$)

Some results (subjective measures)

- Surpisingly, **affect** gets higher ratings on being annoying (5.327 vs. 3.789, p=.0030)
- There was also a 2-way Autonomy*Condition interaction (p=.0058), and the three-way interaction (p=.0358) depicted here
  - affect is viewed as more annoying in two conditions: the **no-autonomy/remote** condition and in the **autonomy/local** condition compared to the other conditions



Implications

- Obtained finer-grained analyses of the effects of robot affect based on attention switches showing that affect has a more complex role than merely as a reminder that time is running out (Donahue and Scheutz, 2015)
- Affect expressions, when paired with simulated robots, negatively affect communication with the robot, while they improve communication when paired with physically embodied robots
- Furthermore, it was shown that female and male participants respond in opposite ways to robot expressions of affect, with females more likely to change attention to the robot when its speech is modulated with affect
- This shows that affect expressions must not be applied in a one-size-fits-all manner, but rather carefully utilized when the aspects of the given scenario are a good fit

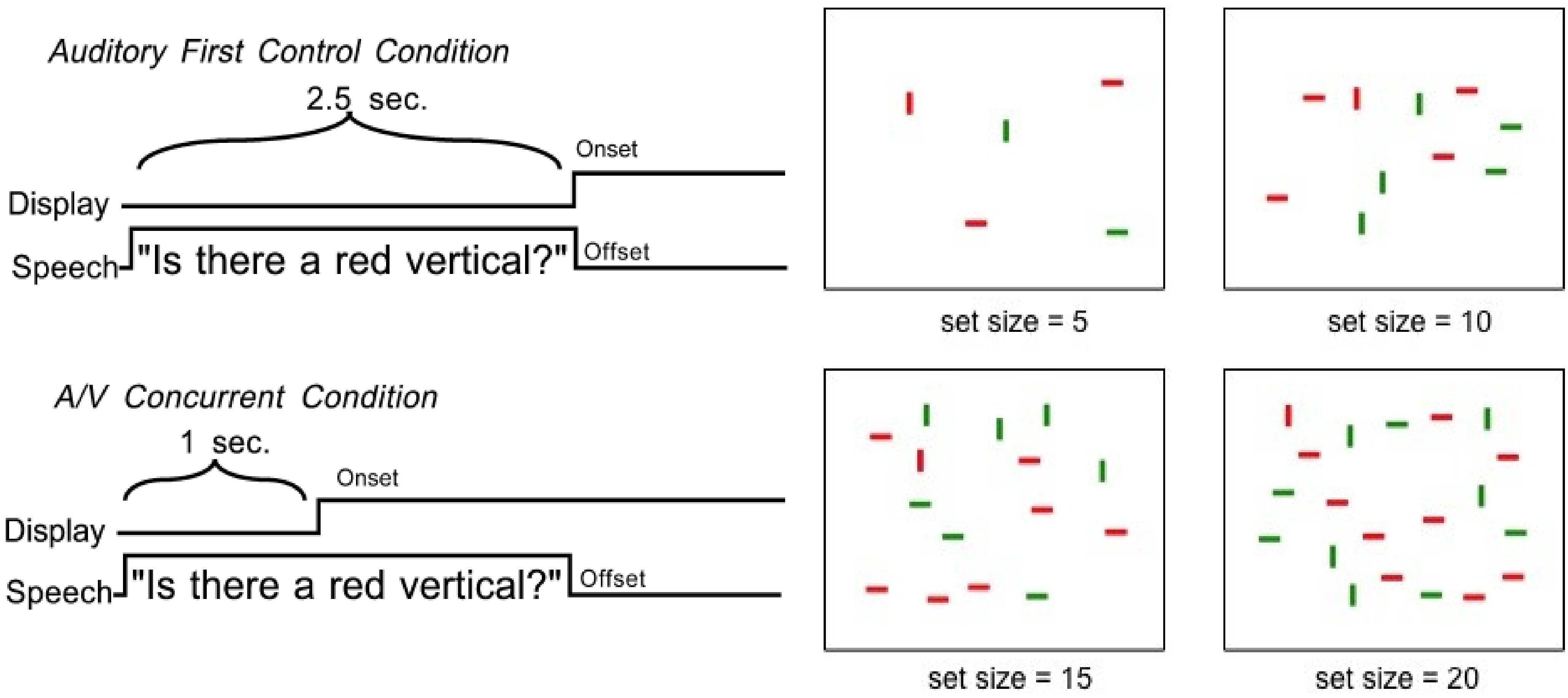
Example 3: conjunctive visual search

Robots have a dual role in cognitive science as

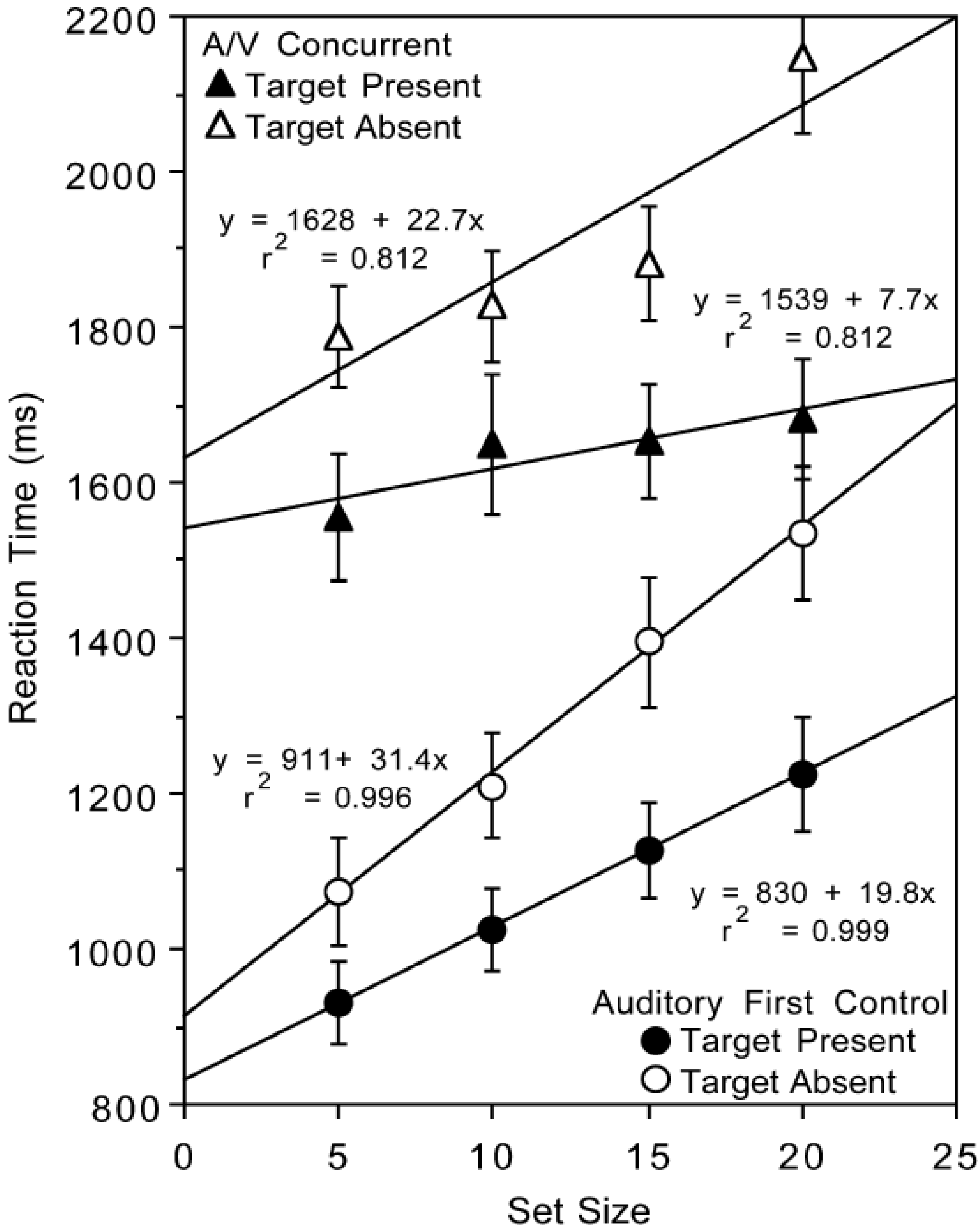
- **tools for studying situated human cognition**
  - *non-cognitive robots* (e.g., robots acting in an environment to study human attention shifts)
  - *cognitive robots* (e.g., the roles of affect and embodiment in task-based team interactions)
- **embodied models of cognition**
  - *non-interaction models* (e.g., models of conjunctive visual search guided by spoken instructions)
  - *interaction models* (e.g., models of indirect speech acts in dialogue-based interactions)

Language guided conjunctive visual search (Spivey et al. 2001)

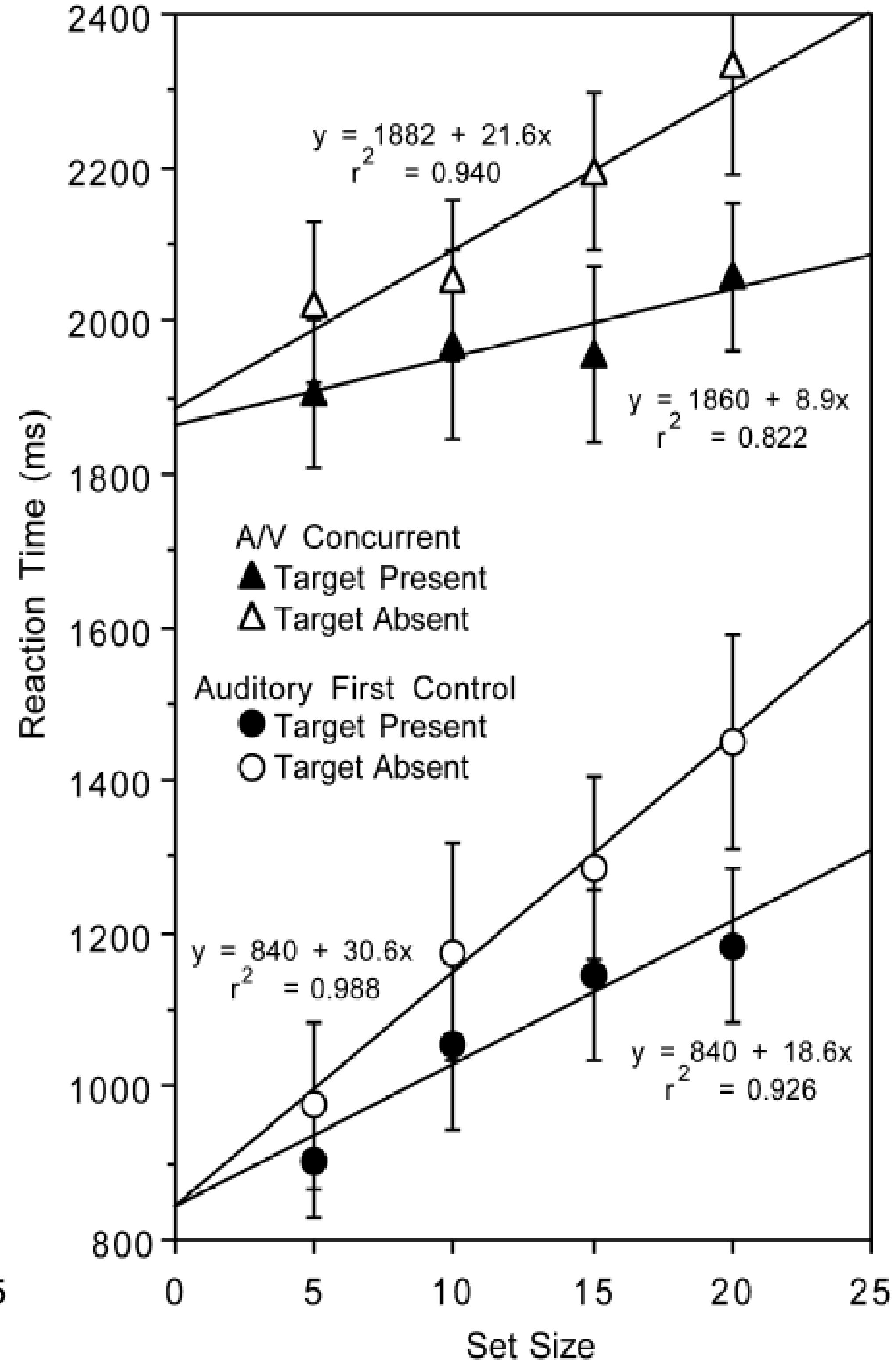
- Vision and natural language processing in humans seem to be highly **interactive** and **incremental**, able to utilize the other modality to reduce processing effort and improve processing (e.g., Eberhard et al. 1995).



Results:

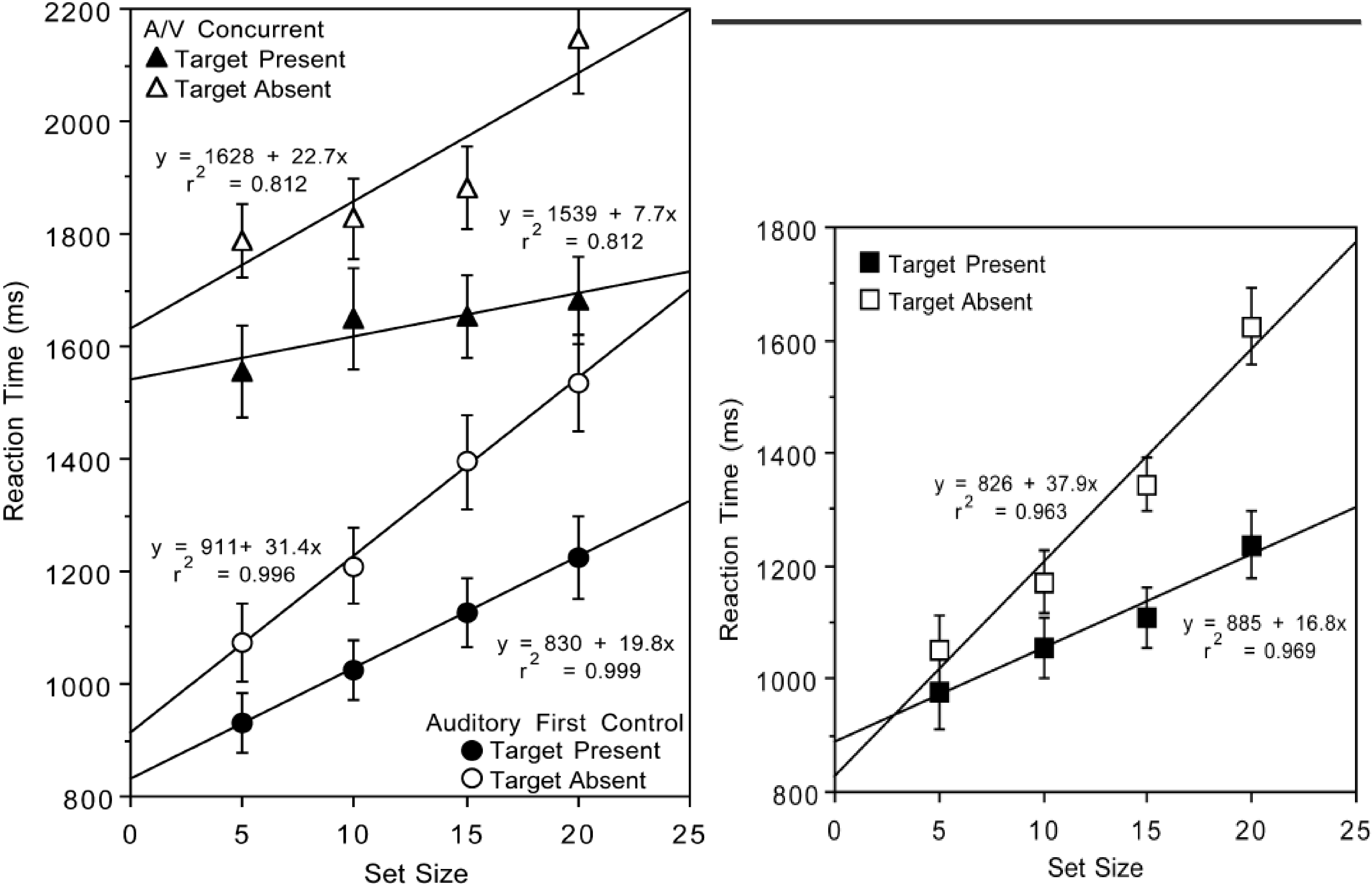


A/V Concurrent
Target Present
Target Absent
y = 1628 + 22.7x
r2 = 0.812
y = 1539 + 7.7x
r2 = 0.812
y = 911+ 31.4x
r2 = 0.996
y = 830 + 19.8x
r2 = 0.999
Auditory First Control
Target Present
Target Absent
Reaction Time (ms)
Set Size
2200
2000
1800
1600
1400
1200
1000
800
0
5
10
15
20
25



y = 1882 + 21.6x
r2 = 0.940
y = 1860 + 8.9x
r2 = 0.822
A/V Concurrent
Target Present
Target Absent
Auditory First Control
Target Present
Target Absent
y = 840 + 30.6x
r2 = 0.988
y = 840 + 18.6x
r2 = 0.926
Reaction Time (ms)
Set Size
2400
2200
2000
1800
1600
1400
1200
1000
800
0
5
10
15
20
25

Results
(compared to visual control)



Conclusions by Spivey et al. 2001

*"In the auditory-first condition, the search process may employ a conjunction template to find the target, thus forcing a **serial-like process** akin to sequentially comparing each object with the target template.*

*However, in the A/V-concurrent condition, it appears that the incremental nature of the speech input allows the search process to begin when only a single feature of the target identity has been heard [which then] **proceeds in a more parallel fashion** (with the second mentioned target feature being used to find the target amidst an attended subset)."*
(Spivey at al. 2001) [emphases are mine]

- Our modeling goal is to evaluate this hypothesis
- Will start by distinguishing **processing mode** (serial vs. parallel) from information **integration mode** (incremental vs non-incremental)

Incremental visual biasing (Krause et al. 2013)

- Use natural language to constrain visual search and focus visual attention, e.g.,

  “Is there a

Incremental visual biasing (Krause et al. 2013)

- Use natural language to constrain visual search and focus visual attention, e.g.,

  “Is there a **tall**

Incremental visual biasing (Krause et al. 2013)

- Use natural language to constrain visual search and focus visual attention, e.g.,

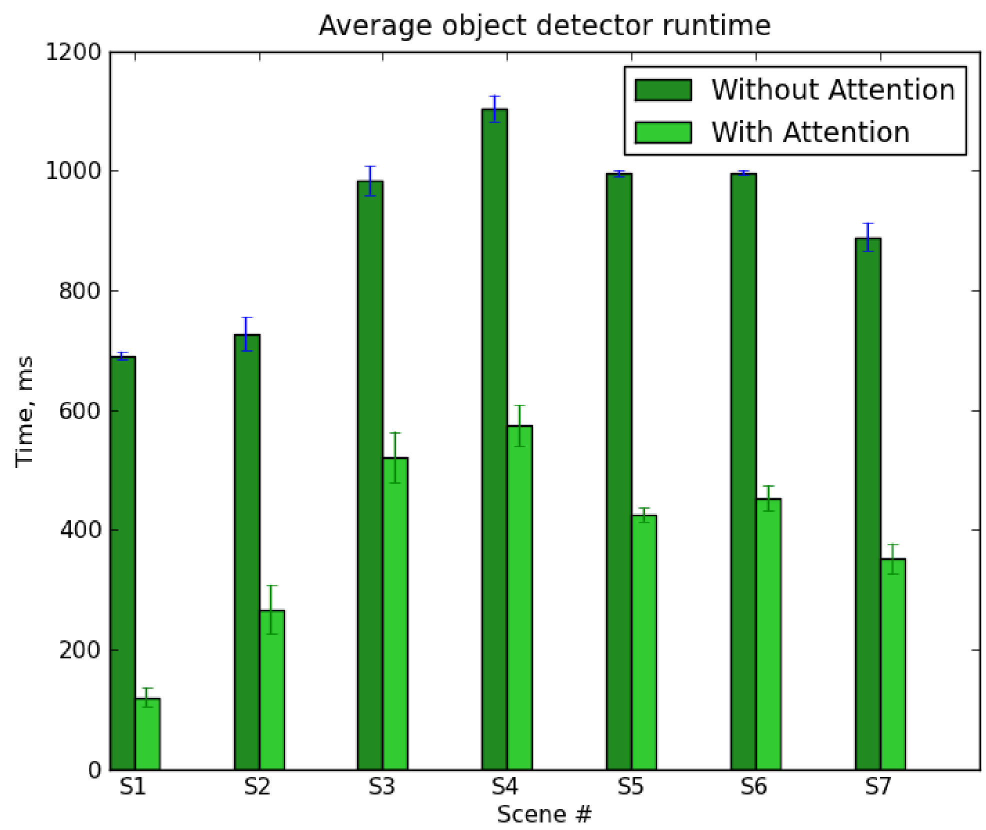
  “Is there a **tall red**

Incremental visual biasing (Krause et al. 2013)

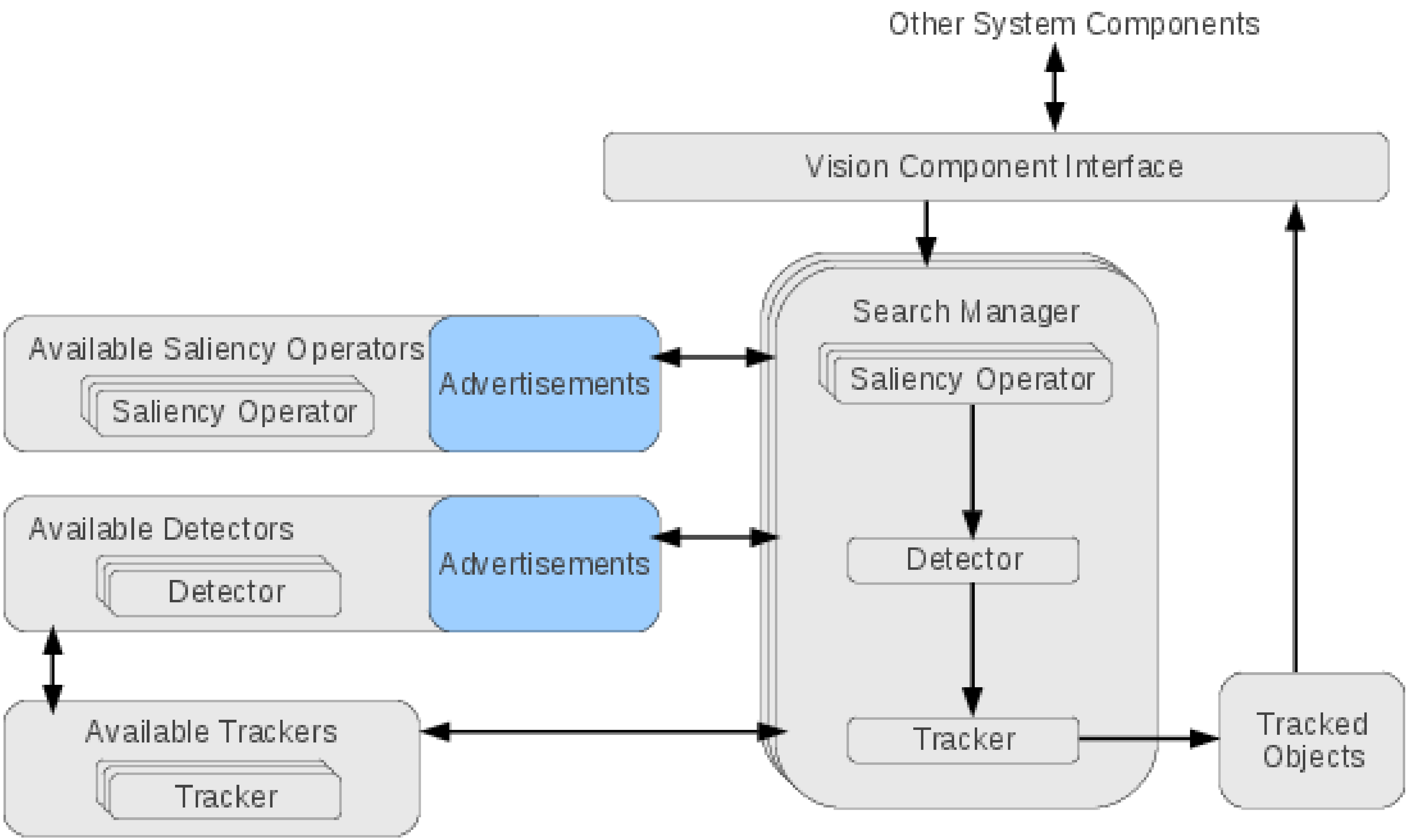
- Use natural language to constrain visual search and focus visual attention, e.g.,

  "Is there a **tall red** object on the **left**?"

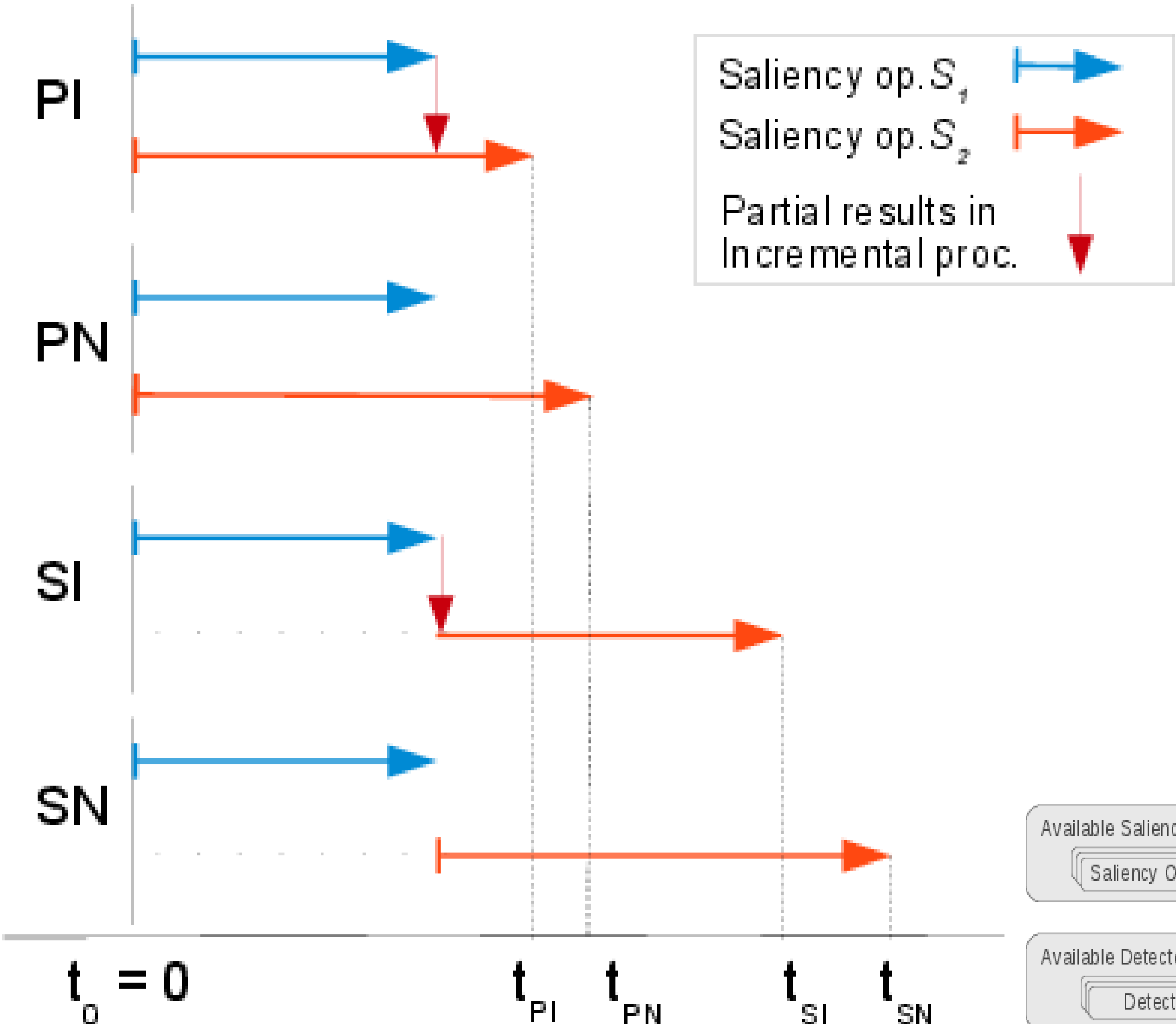
- Improves object detection and reduces computational load

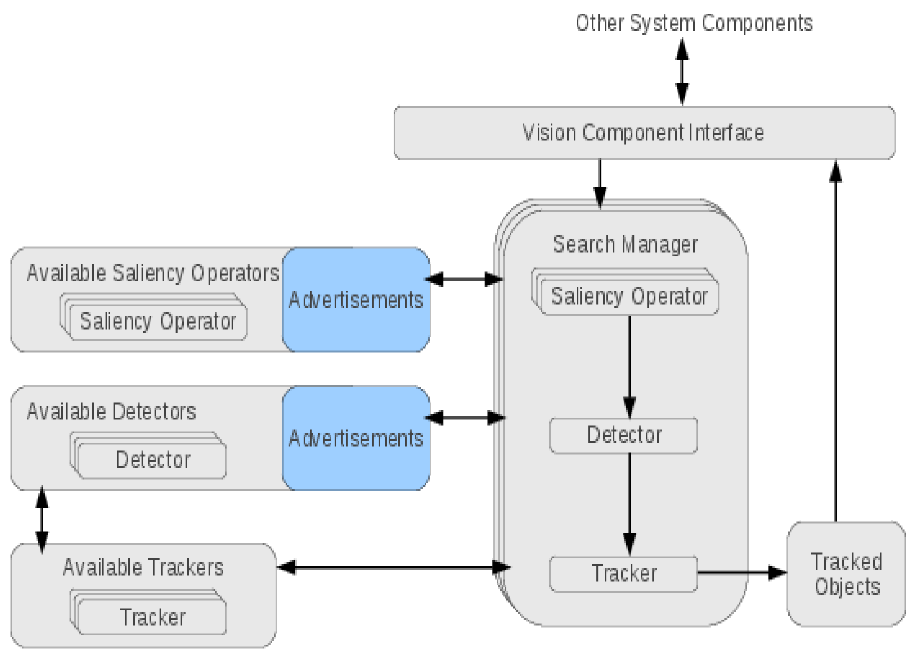


The DIARC Vision System (Krause et al. 2013)



Four configurations of saliency operators (Krause et al. 2013)





Simulations experiments

- We ran 100 replications of each visual stimulus with color terms followed by orientation terms for the four combinations of color (“red” vs. “green”) and orientation (“vertical” vs. “horizontal”) for a total of $100 \cdot 32 \cdot 4 =$ 12800 runs for each of the four model configurations (i.e., over 50000 runs total) in the “audio-first” (A1st) condition
- For each run, we measured the **processing duration for each features** (color and orientation) as well as the **time required for information integration and decision-making** in the object detector
- The resulting output was either “target found” or “target not found” with 100% accuracy

Simulation results

- We performed a 2x2x2x4 ANOVA with **target condition** (*present* vs. *absent*), **integration mode** (*incremental* vs. *non-incremental*), **processing mode** (*parallel* vs. *serial*), and **item set size** (5, 10, 15, or 20 items) as independent, and **total time** (to processing completion from visual stimulus onset) as dependent variables.

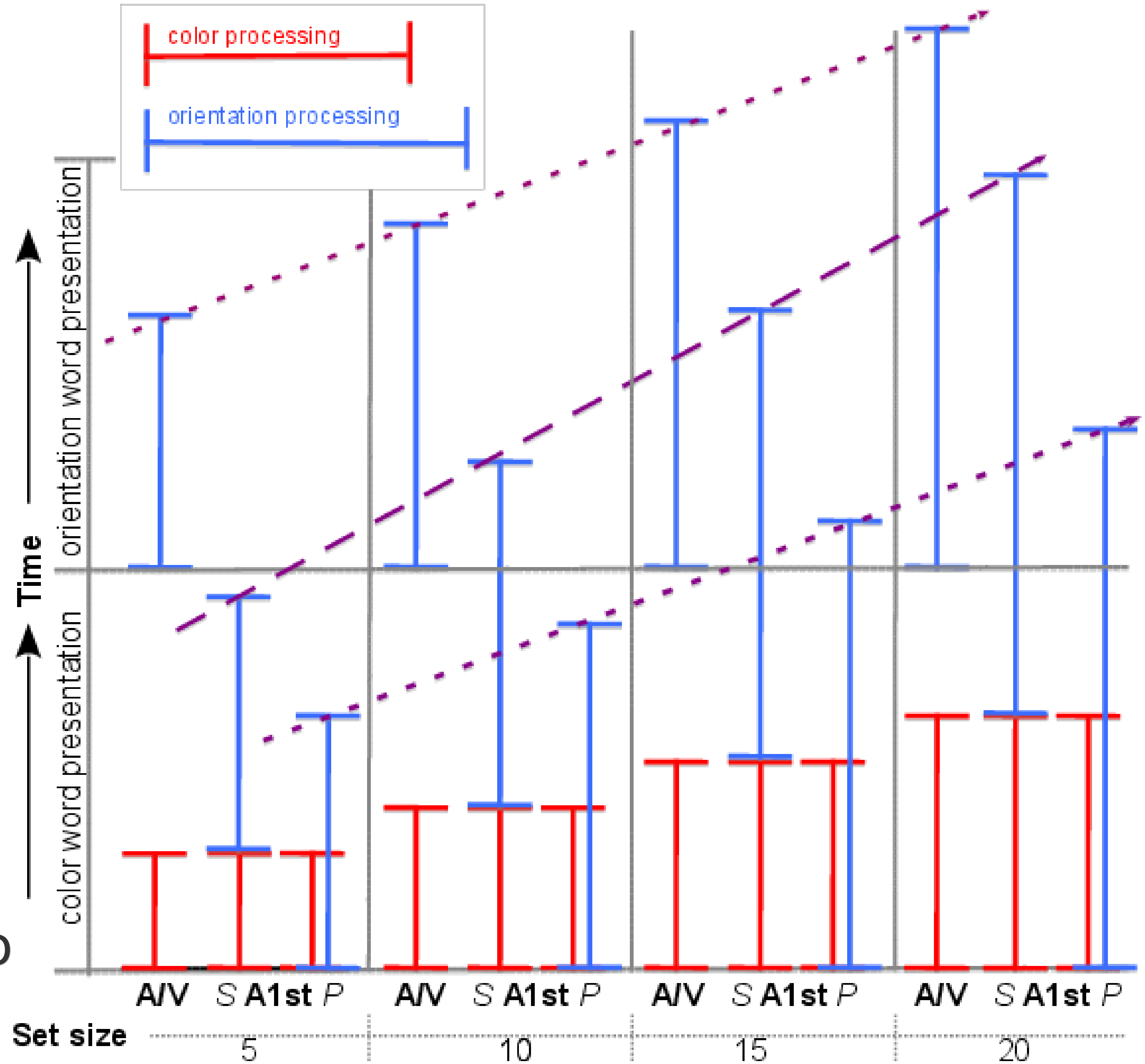
- We found highly significant main effects ($p<.001$) on all four independent variables leading to longer Rts as expected:
  - absence of the target
  - non-incremental processing
  - serial processing
  - increase in item set size

- Compare human **S**(h,p)/**S**(h,a) to model **S**(m,p)/**S**(m,a):

Instr./ Mode	A1st incr.	A1st non-incr.	A/V incr.	A/V non-incr.
serial	2.492462	40.87944	4.633394	75.99336
parallel	4.036554	36.40101	N/A	N/A

Implications

- The model predicts that humans likely use the very same processing configurations in both conditions
- Color processing always completes **before** the onset of the orientation cue (since the cue presentation took 750msec), hence there is no way for the vision system to exploit parallelism



Example 4: indirect speech acts

Robots have a dual role in cognitive science as

- **tools for studying situated human cognition**
  - *non-cognitive robots* (e.g., robots acting in an environment to study human attention shifts)
  - *cognitive robots* (e.g., the roles of affect and embodiment in task-based team interactions)
- **embodied models of cognition**
  - *non-interaction models* (e.g., models of conjunctive visual search guided by spoken instructions)
  - *interaction models* (e.g., models of indirect speech acts in dialogue-based interactions)

Understanding indirect speech acts

- An indirect speech act is an utterance whose literal and intended illocutionary meanings do not match such as

  "Could you get me a coffee?" → ("Get me a coffee")

  "Do you know what time it is?" → ("Tell me the time")

  "Where is the workshop 6?" "Follow me." → ("You will know after you follow my instructions")
- Why do people use indirect speech acts (ISAs)?
- Human social interactions are modulated by many implicit obligations in conjunction with knowledge of social roles and relationships
- *Politeness theory* discusses how indirectness is a strategy for minimizing perceived imposition depending on context and social distance and power relations (e.g., "Give me a coffee!" vs. "Would you be able to get me a coffee?")

Understanding indirect speech acts

- Indirect *requests* have been extensively examined linguistically (e.g., Searle 1975; Asher and Lascarides, 2001) and some computational approaches have been explored in the past (Perrault and Allen, 1980; Wilske and Kruijff, 2006)
- Note that indirect requests are often highly *conventionalized* or *idiomatic* and can be handled by considering

  particular surface form + context →possible indirect request
- We can thus define pragmatic context-dependent rules to handle conventionalized forms (i.e., utterance forms in context together with their pragmatic implication)
- For non-conventionalist forms we can use plan reasoning to infer possible intented meanings

ISA understanding in task-based dialogues

CommX: Do you have a name?
Robot: Yes. My name is Shafer.
CommX: Great. Please stand up.
Robot: Okay.
CommX: Can you to turn right.
Robot: Yes.
CommX: Now stop.
Robot: OK.
CommX: Do you see an obstacle?
Robot: Yes.
CommX: Walk forward.
Robot: I am sorry but I cannot do that because there is an obstacle ahead.
CommX: The obstacle is not solid.
Robot: Okay.
CommX: Move forward.
Robot: Okay.
Robot: I am sorry but I cannot do that because there is no support ahead.
CommX: Okay. Just go.
Robot: But it is not safe.
CommX: I will catch you.
Robot: Okay.
CommX: Now I need you to go straight.
Robot: Okay.

ISA understanding in task-based dialogues

CommX: Do you have a name?
Robot: Yes. My name is Shafer.
CommX: Great. Please stand up.
Robot: Okay.
CommX: Can you to turn right.
Robot: Yes.
CommX: Now stop.
Robot: OK.
CommX: Do you see an obstacle?
Robot: Yes.
CommX: Walk forward.
Robot: I am sorry but I cannot do that because there is an obstacle ahead.
CommX: The obstacle is not solid.
Robot: Okay.
CommX: Move forward.
Robot: Okay.
Robot: I am sorry but I cannot do that because there is no support ahead.
CommX: Okay. Just go.
Robot: But it is not safe.
CommX: I will catch you.
Robot: Okay.
CommX: Now I need you to go straight.
Robot: Okay.

ISA understanding in task-based dialogues

CommX: Do you have a name?
Robot: Yes. My name is Shafer.
CommX: Great. Please stand up.
Robot: Okay.
CommX: Can you to turn right.
Robot: Yes.
CommX: Now stop.
Robot: OK.
CommX: Do you see an obstacle?
Robot: Yes.
CommX: Walk forward.
Robot: I am sorry but I cannot do that because there is an obstacle ahead.
CommX: The obstacle is not solid.
Robot: Okay.
CommX: Move forward.
Robot: Okay.
Robot: I am sorry but I cannot do that because there is no support ahead.
CommX: Okay. Just go.
Robot: But it is not safe.
CommX: I will catch you.
Robot: Okay.
CommX: Now I need you to go straight.
Robot: Okay.

Pragmatic Reasoning

AskYN(commX,self,have(self,name),{})

ISA understanding in task-based dialogues

CommX: Do you have a name?
Robot: Yes. My name is Shafer.
CommX: Great. Please stand up.
Robot: Okay.
CommX: Can you to turn right.
Robot: Yes.
CommX: Now stop.
Robot: OK.
CommX: Do you see an obstacle?
Robot: Yes.
CommX: Walk forward.
Robot: I am sorry but I cannot do that because there is an obstacle ahead.
CommX: The obstacle is not solid.
Robot: Okay.
CommX: Move forward.
Robot: Okay.
Robot: I am sorry but I cannot do that because there is no support ahead.
CommX: Okay. Just go.
Robot: But it is not safe.
CommX: I will catch you.
Robot: Okay.
CommX: Now I need you to go straight.
Robot: Okay.

Pragmatic Reasoning

AskYN(commX,self,have(self,name),{})

↓

Literal Interpretation

bel(self,itk(commX,have(self,name)))

Non-Literal Interpretation

bel(self,itkRef(commX,nameOf(self)))

ISA understanding in task-based dialogues

CommX: Do you have a name?
Robot: Yes. My name is Shafer.
CommX: Great. Please stand up.
Robot: Okay.
CommX: Can you to turn right.
Robot: Yes.
CommX: Now stop.
Robot: OK.
CommX: Do you see an obstacle?
Robot: Yes.
CommX: Walk forward.
Robot: I am sorry but I cannot do that because there is an obstacle ahead.
CommX: The obstacle is not solid.
Robot: Okay.
CommX: Move forward.
Robot: Okay.
Robot: I am sorry but I cannot do that because there is no support ahead.
CommX: Okay. Just go.
Robot: But it is not safe.
CommX: I will catch you.
Robot: Okay.
CommX: Now I need you to go straight.
Robot: Okay.

Pragmatic Reasoning

AskYN(commX,self,have(self,name),{})

Literal Interpretation

bel(self,itk(commX,have(self,name)))

Non-Literal Interpretation

bel(self,itkRef(commX,nameOf(self)))

Belief Reasoning/Question Answering

bel(self,itk(commX,have(self,name))) ^
bel(self,have(self,name))

ISA understanding in task-based dialogues

CommX: **Do you have a name?**
Robot: Yes. My name is Shafer.
CommX: Great. Please stand up.
Robot: Okay.
CommX: **Can you to turn right.**
Robot: Yes.
CommX: Now stop.
Robot: OK.
CommX: Do you see an obstacle?
Robot: Yes.
CommX: Walk forward.
Robot: I am sorry but I cannot do that because there is an obstacle ahead.
CommX: The obstacle is not solid.
Robot: Okay.
CommX: Move forward.
Robot: Okay.
Robot: I am sorry but I cannot do that because there is no support ahead.
CommX: Okay. Just go.
Robot: But it is not safe.
CommX: I will catch you.
Robot: Okay.
CommX: **Now I need you to go straight.**
Robot: Okay.

Pragmatic Reasoning

AskYN(commX,self,have(self,name),{})

Literal Interpretation

bel(self,itk(commX,have(self,name)))

Non-Literal Interpretation

bel(self,itkRef(commX,nameOf(self)))

Belief Reasoning/Question Answering

bel(self,itk(commX,have(self,name))) ^
bel(self,have(self,name))

Generate Reply: “Yes.”

ISA understanding in task-based dialogues

CommX: Do you have a name?
Robot: Yes. My name is Shafer.
CommX: Great. Please stand up.
Robot: Okay.
CommX: Can you to turn right.
Robot: Yes.
CommX: Now stop.
Robot: OK.
CommX: Do you see an obstacle?
Robot: Yes.
CommX: Walk forward.
Robot: I am sorry but I cannot do that because there is an obstacle ahead.
CommX: The obstacle is not solid.
Robot: Okay.
CommX: Move forward.
Robot: Okay.
Robot: I am sorry but I cannot do that because there is no support ahead.
CommX: Okay. Just go.
Robot: But it is not safe.
CommX: I will catch you.
Robot: Okay.
CommX: Now I need you to go straight.
Robot: Okay.

Pragmatic Reasoning

AskYN(commX,self,have(self,name),{})

Literal Interpretation

bel(self,itk(commX,have(self,name)))

Non-Literal Interpretation

bel(self,itkRef(commX,nameOf(self)))

Belief Reasoning/Question Answering

bel(self,itkRef(commX,nameOf(self))) ^
bel(self,is(nameOf(self),Shafer))

Generate Statement: “My name is Shafer.”

ISA understanding in task-based dialogues

CommX: Do you have a name?
Robot: Yes. My name is Shafer.
CommX: Great. Please stand up.
Robot: Okay.
CommX: Can you to turn right.
Robot: Yes.
CommX: Now stop.
Robot: OK.
CommX: Do you see an obstacle?
Robot: Yes.
CommX: Walk forward.
Robot: I am sorry but I cannot do that because there is an obstacle ahead.
CommX: The obstacle is not solid.
Robot: Okay.
CommX: Move forward.
Robot: Okay.
Robot: I am sorry but I cannot do that because there is no support ahead.
CommX: Okay. Just go.
Robot: But it is not safe.
CommX: I will catch you.
Robot: Okay.
CommX: Now I need you to go straight.
Robot: Okay.

Pragmatic Reasoning

AskYN(commX,self,capableOf(self,turn(self,right)),{})

ISA understanding in task-based dialogues

CommX: Do you have a name?
Robot: Yes. My name is Shafer.
CommX: Great. Please stand up.
Robot: Okay.
CommX: Can you to turn right.
Robot: Yes.
CommX: Now stop.
Robot: OK.
CommX: Do you see an obstacle?
Robot: Yes.
CommX: Walk forward.
Robot: I am sorry but I cannot do that because there is an obstacle ahead.
CommX: The obstacle is not solid.
Robot: Okay.
CommX: Move forward.
Robot: Okay.
Robot: I am sorry but I cannot do that because there is no support ahead.
CommX: Okay. Just go.
Robot: But it is not safe.
CommX: I will catch you.
Robot: Okay.
CommX: Now I need you to go straight.
Robot: Okay.

Pragmatic Reasoning

AskYN(commX,self,capableOf(self,turn(self,right)),{})

Literal Interpretation

bel(self,itk(commX,capableOf(self,turning(self,right))))

Non-Literal Interpretation

bel(self,want(commX,turning(self,right)))

ISA understanding in task-based dialogues

CommX: Do you have a name?
Robot: Yes. My name is Shafer.
CommX: Great. Please stand up.
Robot: Okay.
CommX: Can you to turn right.
Robot: Yes.
CommX: Now stop.
Robot: OK.
CommX: Do you see an obstacle?
Robot: Yes.
CommX: Walk forward.
Robot: I am sorry but I cannot do that because there is an obstacle ahead.
CommX: The obstacle is not solid.
Robot: Okay.
CommX: Move forward.
Robot: Okay.
Robot: I am sorry but I cannot do that because there is no support ahead.
CommX: Okay. Just go.
Robot: But it is not safe.
CommX: I will catch you.
Robot: Okay.
CommX: Now I need you to go straight.
Robot: Okay.

Pragmatic Reasoning

AskYN(commX,self,capableOf(self,turn(self,right)),{})

Literal Interpretation

bel(self,itk(commX,capableOf(self,turning(self,right))))

Non-Literal Interpretation

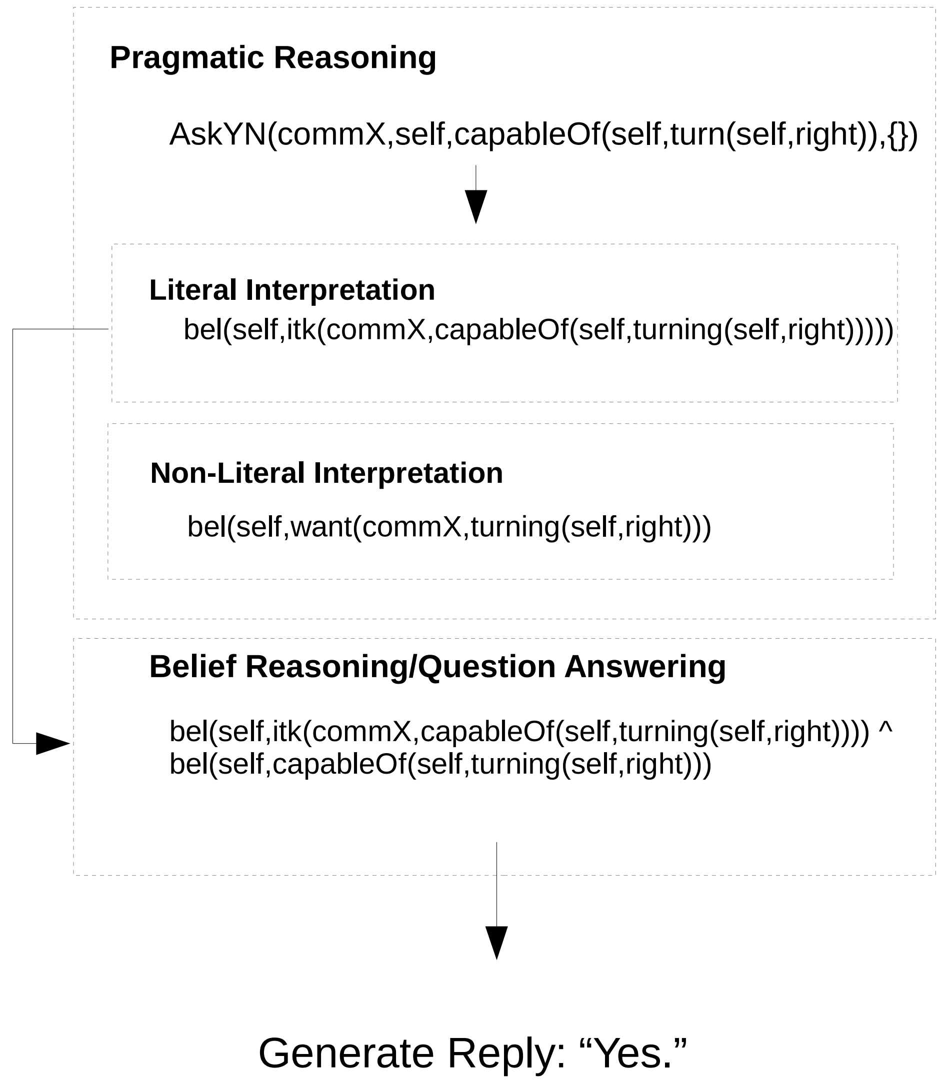
bel(self,want(commX,turning(self,right)))

Belief Reasoning/Question Answering

bel(self,itk(commX,capableOf(self,turning(self,right)))) ^
bel(self,capableOf(self,turning(self,right)))

ISA understanding in task-based dialogues

CommX: Do you have a name?
Robot: Yes. My name is Shafer.
CommX: Great. Please stand up.
Robot: Okay.
CommX: Can you to turn right.
Robot: Yes.
CommX: Now stop.
Robot: OK.
CommX: Do you see an obstacle?
Robot: Yes.
CommX: Walk forward.
Robot: I am sorry but I cannot do that because there is an obstacle ahead.
CommX: The obstacle is not solid.
Robot: Okay.
CommX: Move forward.
Robot: Okay.
Robot: I am sorry but I cannot do that because there is no support ahead.
CommX: Okay. Just go.
Robot: But it is not safe.
CommX: I will catch you.
Robot: Okay.
CommX: Now I need you to go straight.
Robot: Okay.



ISA understanding in task-based dialogues

CommX: Do you have a name?
Robot: Yes. My name is Shafer.
CommX: Great. Please stand up.
Robot: Okay.
CommX: Can you to turn right.
Robot: Yes.
CommX: Now stop.
Robot: OK.
CommX: Do you see an obstacle?
Robot: Yes.
CommX: Walk forward.
Robot: I am sorry but I cannot do that because there is an obstacle ahead.
CommX: The obstacle is not solid.
Robot: Okay.
CommX: Move forward.
Robot: Okay.
Robot: I am sorry but I cannot do that because there is no support ahead.
CommX: Okay. Just go.
Robot: But it is not safe.
CommX: I will catch you.
Robot: Okay.
CommX: Now I need you to go straight.
Robot: Okay.

Pragmatic Reasoning

AskYN(commX,self,capableOf(self,turn(self,right)),{})

Literal Interpretation

bel(self,itk(commX,capableOf(self,turning(self,right))))

Non-Literal Interpretation

bel(self,want(commX,turning(self,right)))

Goal Adoption Reasoning

obl(self,turning(self,right)) ^
not(per(self,not(turning(self,right))))
→ goal(self,turning(self,right))

Robot beings to turn right.

Implications

- Provided an architecture for understanding simple conventionalized ISAs in task-based dialogues
- Need to determine the extent to which people will likely use them in a given task (which will depend on people's familiarity with the task, etc. – preliminary evidence from HRI experiments suggests that it is difficult for people to not use ISAs in conventionalized tasks)
- Need a systematic way of handling ISAs that are not conventionalized (e.g., using intent recognition, plan reasoning, etc.)
- Use model to generate ISAs and to predict when people will use ISAs

Conclusions

- We argued that robots can serve a **dual role** in cognitive science, both as experimental tools and embodied models
- Human-robot interaction is an ideal area that brings together both uses in order to better understand human cognition as well as building technologies for future robots
- Demonstrated both roles with four examples:
  - robots used as tools to study joint attention as well as the interaction of affect and embodiment, and
  - robots used as models to model human language-guided conjunctive visual search and indirect speech act understanding in task-based dialogues
- With increasing sophistication of AI and robot technology, we will eventually be able to build more sophisticated tools and models to study human cognition and interactions in realistic multi-agent tasks and environments

Acknowledgments and links

- Former students/postdocs: Gordon Briggs, Rehj Cantrell, Matt Dunlap, Paul Schermerhorn, Richard Veale, Tom Donahue
- Current students/staff: Evan Krause, Tom Williams, Brad Oosteveld, Zach Haga, and various undergraduate students at Indiana University and Tufts University
- Collaborators: Chen Yu's group at Indiana University (joint attention)
- Funding provided by the US National Science Foundation and the US Office of Naval Research through several grants

- Papers at http://hrilab.tufts.edu/publications/
- Demos at https://www.youtube.com/user/HRILaboratory