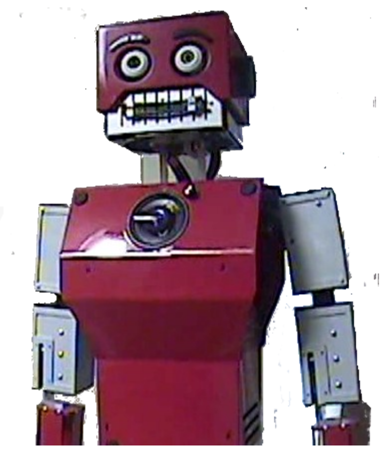
Robots in Cognition and Cognition in Robots:
The Dual Role of Robots in Cognitive Science

Matthias Scheutz
mscheutz@cs.tufts.edu
Director, Human-Robot Interaction Laboratory
Director, Cognitive Science Ph.D. Program
Department of Computer Science
Department of Psychology
Tufts University
Medford, MA 02155
http://hrilab.cs.tufts.edu/

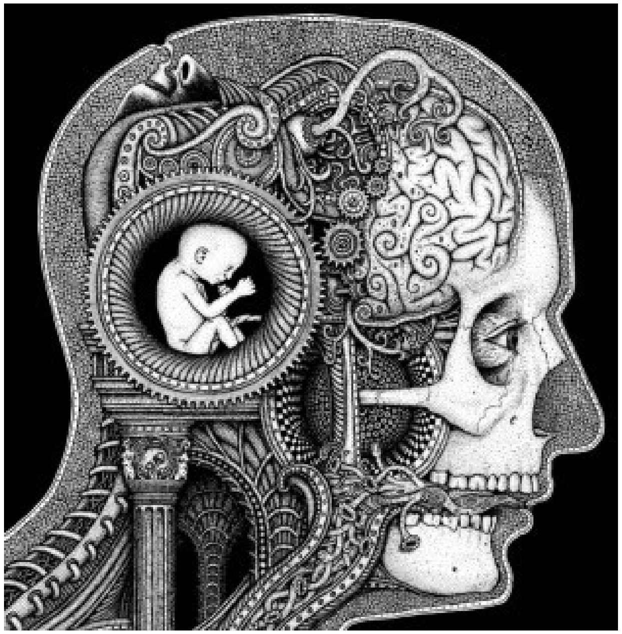


Computationalism and "Classical" Robotics

- Cognitive science was founded on the "computer metaphor":

Mind:Brain ~ Program:Computer

- **Computationalism** then is the view that mental processes *are* computational processes, and consequently, that cognitive functions can be fully described by programs
- Related, but different from the computational claim is what Searle called the "**homunculus fallacy**", namely the view of treating "the brain as if there were some agent inside it using it to compute with" (cp. to Dennett's "Cartesian Theater")
- "Classical robotics" adopted the idea of a central controller being in charge of the robot's body, taking in sensory signals, constructing a "model of the world" for planning the actions to be carried out



(Failed) Representations and World(-less) Models

- While in cognitive science, “representations” were used as explanatory vehicles in cognitive computational theories, “representations” in robotic control architectures were used to allow robots to perform operations that did not immediately result in actions (e.g., planning)
- The notion of **representation** was subsequenly challenged in cognitive science (most notably by some connectionists and dynamicists) and robotics (most notably by behavior-based and evolutionary robotics camps): neither human cognition nor robotics control was viewed as performing syntactic operations on abstract representations, albeit for different reasons
- These shifts in perspective brought robots and cognition closer together in the 1990ies than ever before, in what has been dubbed “**situated embodied cognition**”

Focus on the Body

- **Embodied cognitive science** was a reaction to classical (non-embodied) cognitive science that, as Clark 1999 puts it,
  - understanding the complex interplay of brain, body and world requires new analytic tools and methods, such as those of dynamical systems theory
  - traditional notions of internal representation and computation are inadequate and unnecessary
  - the typical decomposition of the cognitive system into a variety of inner neural or functional subsystems is often misleading, and blinds us to the possibility of alternative, and more explanatory, decompositions that cut across the traditional brain–body–world divisions
- And he asks: “If we follow the embodied, embedded approach to its natural conclusions, do we lose sight of the differences between perception, reason and action?”

Focus on the Situation

- **Situated cognitive science** extends and complements the embodied approach to cognition by acknowledging the role situations play in human cognition and action.

  "Human knowledge and interaction cannot be divorced from the world. To do so is to study a disembodied intelligence, one that is artificial, unreal, and uncharacteristic of actual behavior. What really matters is the situation and the parts that people play. One cannot look at just the situation, or just the environment, or just the person. [...] After all, it is the mutual accommodation of people and the environment that matters, so to focus upon only aspects in isolation is to destroy the interaction, to eliminate the role of the situation upon cognition and action." (Norman 1993)

- In particular, social interactions and their dynamics posed new challenges for **studying**, **modeling** and **explaining** cognition

Focus on Logic

- In AI, the need for robots was probably best articulated in the Reiter 1993 IJCAI Award for Research Excellence lecture where he publicly introduced the term “cognitive robotics”
- What is it? Hector Levesque and Ray Reiter wrote in their position paper for the 1998 Cognitive Robotics workshop:

  “We agree with the premise of the workshop that it is time to take seriously the need for high-level cognition in designing robotic systems... For the past five years a group of us at the University of Toronto have been engaged in what we call Cognitive Robotics, which we take to be the study of the knowledge representations and reasoning problems faced by an autonomous robot (or agent) in a dynamic incompletely known world. Central to this effort is to develop an understanding of the relation between the knowledge, the perception, and the action of such an robot.”

“Impliers”

- The bottom lever from proponents of sensory-motor couplings in cognitive science paired with top lever from logic-based AI, which were both motivated to use robots to address shortcomings in their fields, led to a firm hold of robotics in cognitive science (and AI)
- And while low-level and high-level approaches to embodied cognition proceeded largely independently, we are now witnessing a surge of integrated robotic architectures for artificial cognitive systems

2007

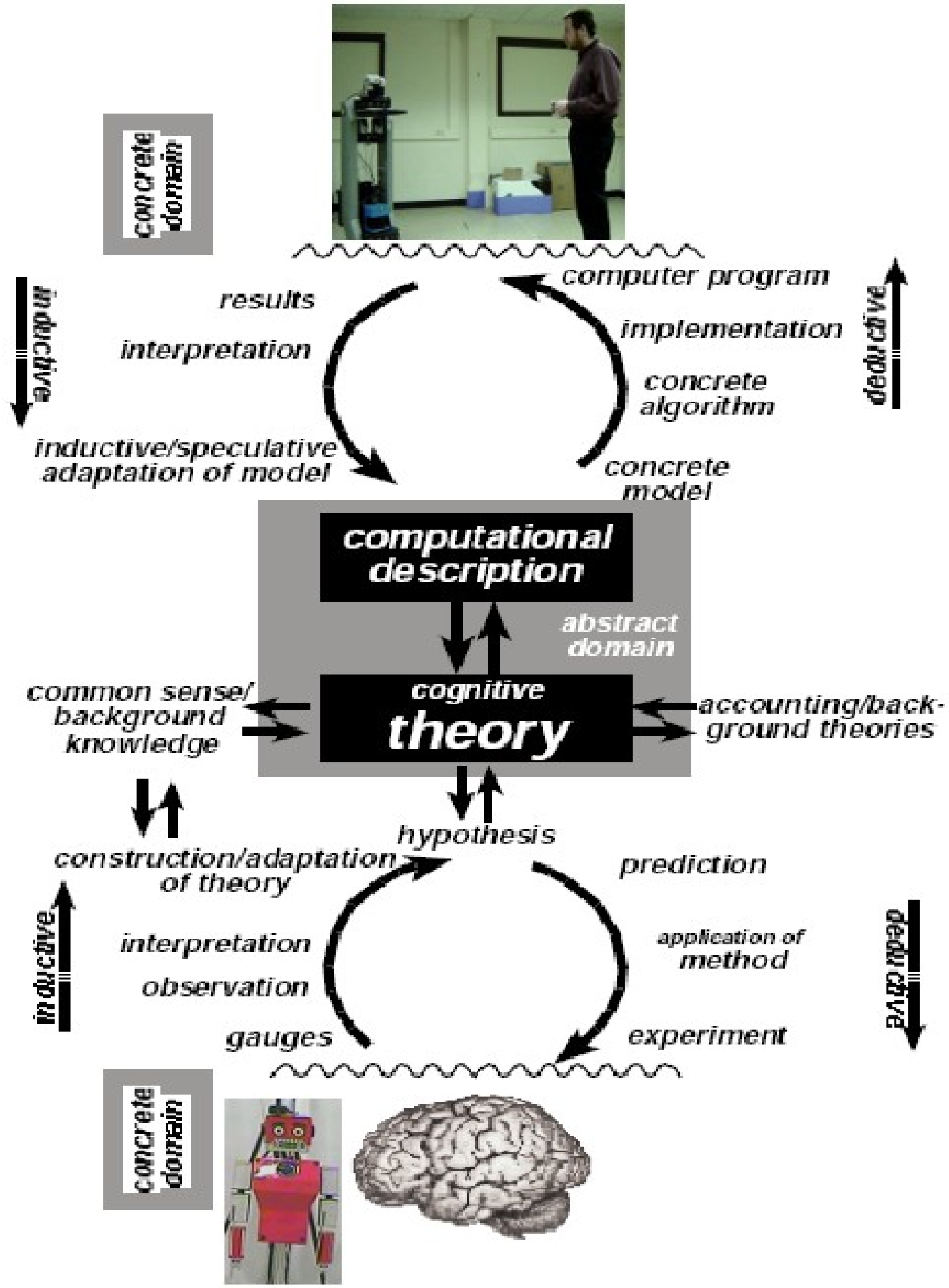
The dual role of robots in cognitive science

Robots as experimental tools

to better study human social cognitive processes that unfold in real-time where responses depend on past behaviors (e.g., allowing for “contingent experimental designs”)

Robots as cognitive models

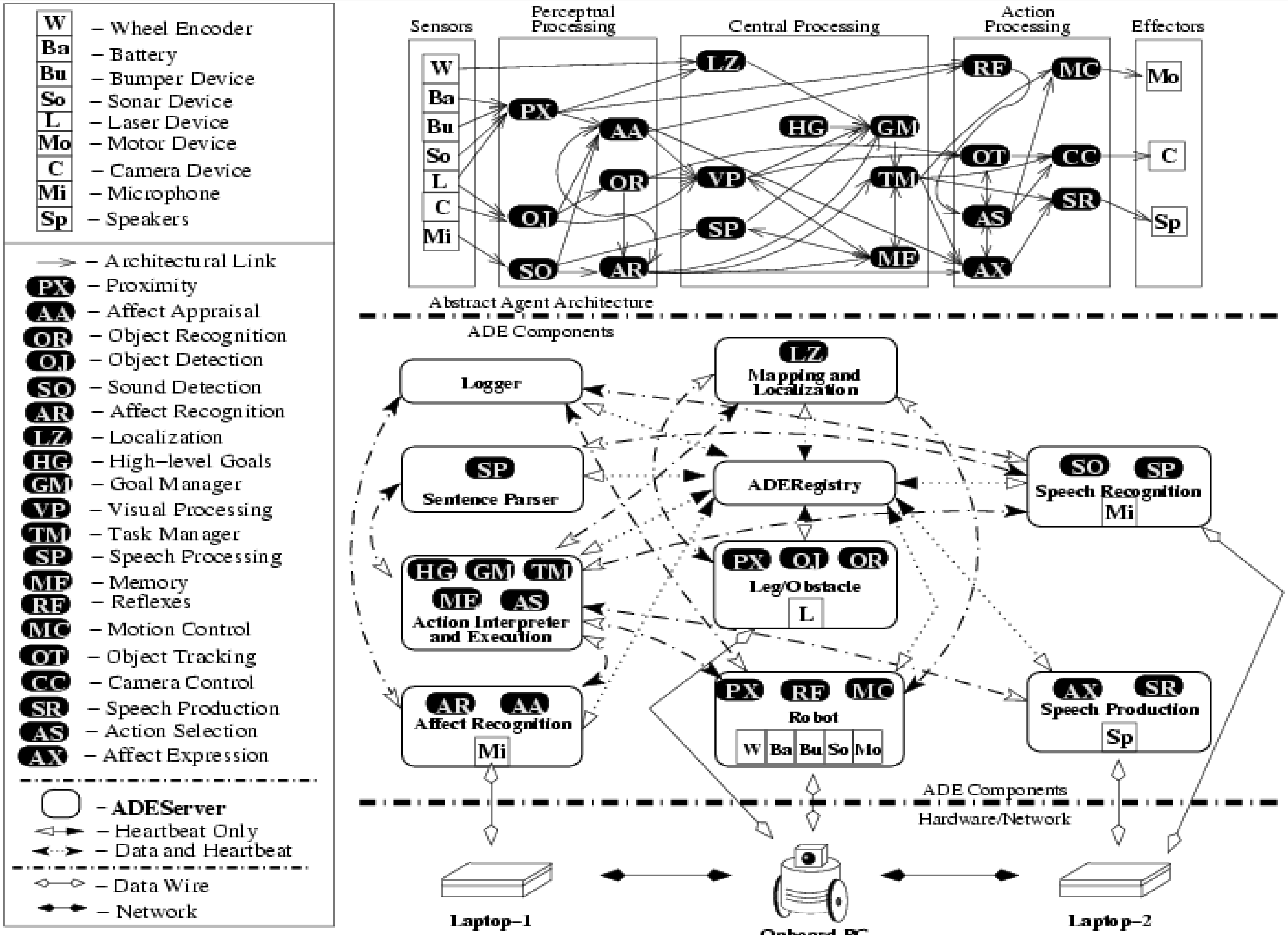
to develop and implement computational models for situated embodied cognition that can also be used for practical applications as they are able to perform tasks in real-world situations



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., Cantrell et al. 2010, Scheutz et al. 2010, Schermerhorn and Scheutz 2010, Scheutz et al., 2007 Schermerhorn et al. 2005, and others)
  - **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)

(Parts of) DIARC implemented in the ADE middleware



DIARC's theoretical commitments

- All processing in components occurs asynchronously to other components (e.g., as in the subsumption architecture)
- Each components operates on a “cognitive cycle” (the “loop time”) and may run multiple threads of control within itself (e.g., perception, natural language processing, and action execution are examples of highly parallelized components)
- Goals are explicitly represented in terms of pre- operating- and post-conditions, have a priority that is computed based on urgency, expected utilities and overal affective state (which is computed for each component based on its operation), and are attached to skills that accomplish them
- Behavior arbitration is distributed and priority-based (using goal priorities) and uses hierarchical locking mechanisms for mutual exlusion of effectors and other architectural resources (e.g., memory access, speech output, etc.)

DIARC's theoretical commitments

- Different forms of learning occur in different components (e.g., statistical learning in components close to perception and action, symbolic learning in higher-level components)
- Knowledge representations take different forms within components depending on the nature of the process (e.g., saliency maps inside the vision processing component, dependency graphs in the parser, clauses in the reasoner,...)
- Particular types of semantic expressions (represented in formal logic) are used as a "common currency" and data representation format across components wherever possible (e.g., between NL and vision) and are used as part of introspective access to system features and capbilities
- Processing "hooks" into the implementation platform allow components to introspect and monitor the operation of parts of the system (and allow for discovery of system features and failures using the ADE notification mechanisms)

DIARC compared to classical cognitive architectures

- Classical cognitive architectures such as **ACT-R** or **SOAR** are typically monolithic and operate sequentially on a cognitive cycle (even though some have limited parallelism, e.g., Epic)
- Sensory processing and effector control are typically very limited and occur at a high-level of abstraction (e.g., visual objects are assumed or action control is often sequential)
- There are limited mechanisms for handling parallelism and the (reference) implementations are typically not distributed
- Neither SOAR nor ACT-R have been used for the control of robots as measurement instruments in HRI studies
- And while both SOAR and ACT-R have been used to model human performance in social contexts, they have typically not been used to implement situated computational models on robots that are evaluated in HRI experiments (although there some ACT-R models were used in HRI studies)

Robots as tools: examples from our work

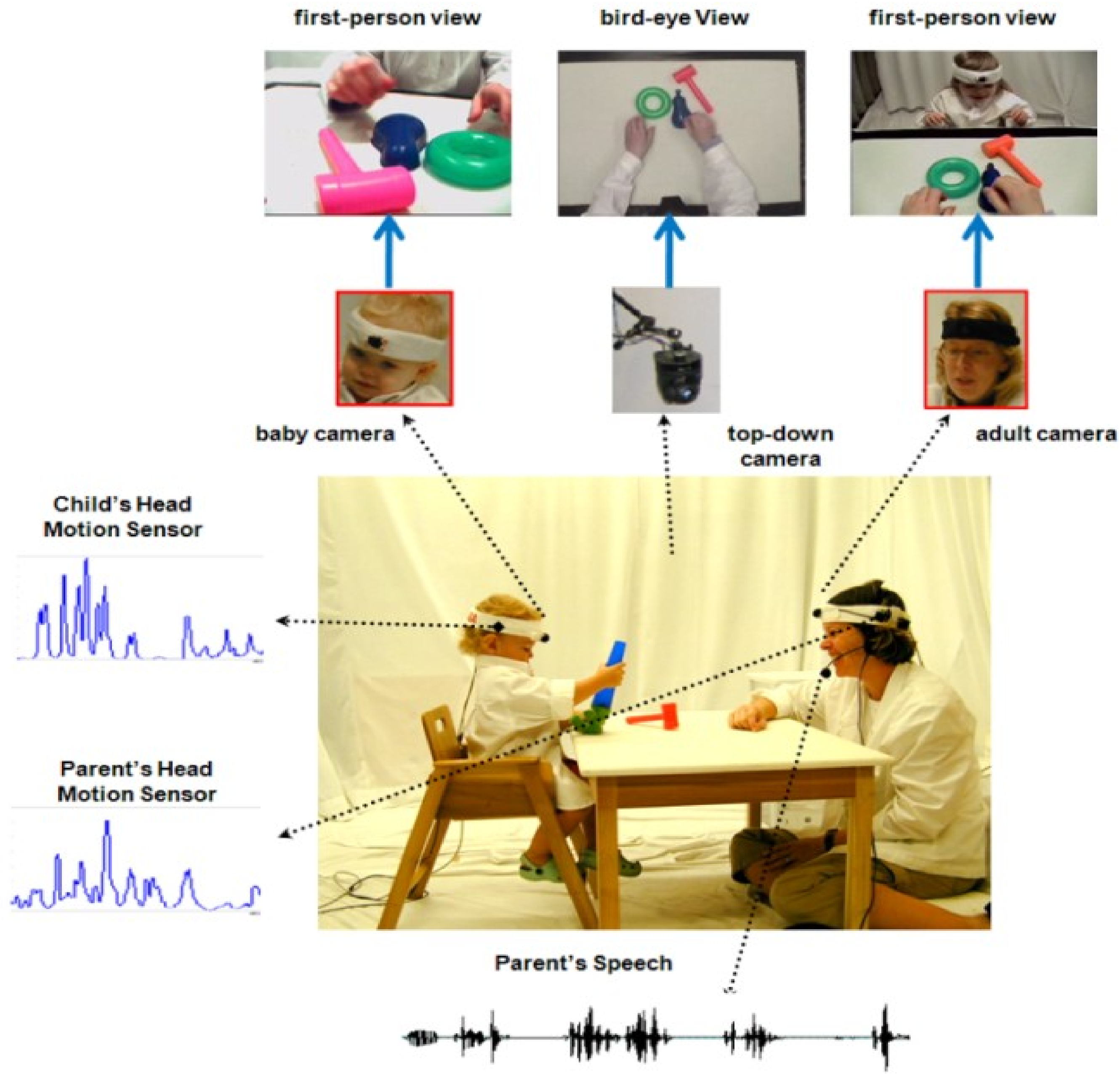
- Joint attention processes:
  - establishing and maintaining joint attention
  - breaking joint attention (through “abnormal attention”)
- Human attitudes about robots:
  - social facilitation and social inhibition (to probe agency)
  - robotic voice, social presence and gender differences
- Human reactions to autonomous robots in cooperative tasks:
  - human reactions to robot affect
  - human reactions to real vs simulated robots w/o autonomy
- Task-switching in human multi-tasking:
  - fNIRs-based adaption of robot autonomy
  - effects of real vs virtual robots on multi-tasking performance
- Philosophical and conceptual inquiry:
  - what it is like to be an agent/have a red experience?
  - the effects of “ethical robots” on human decision-making

Robots as tools: examples from our work

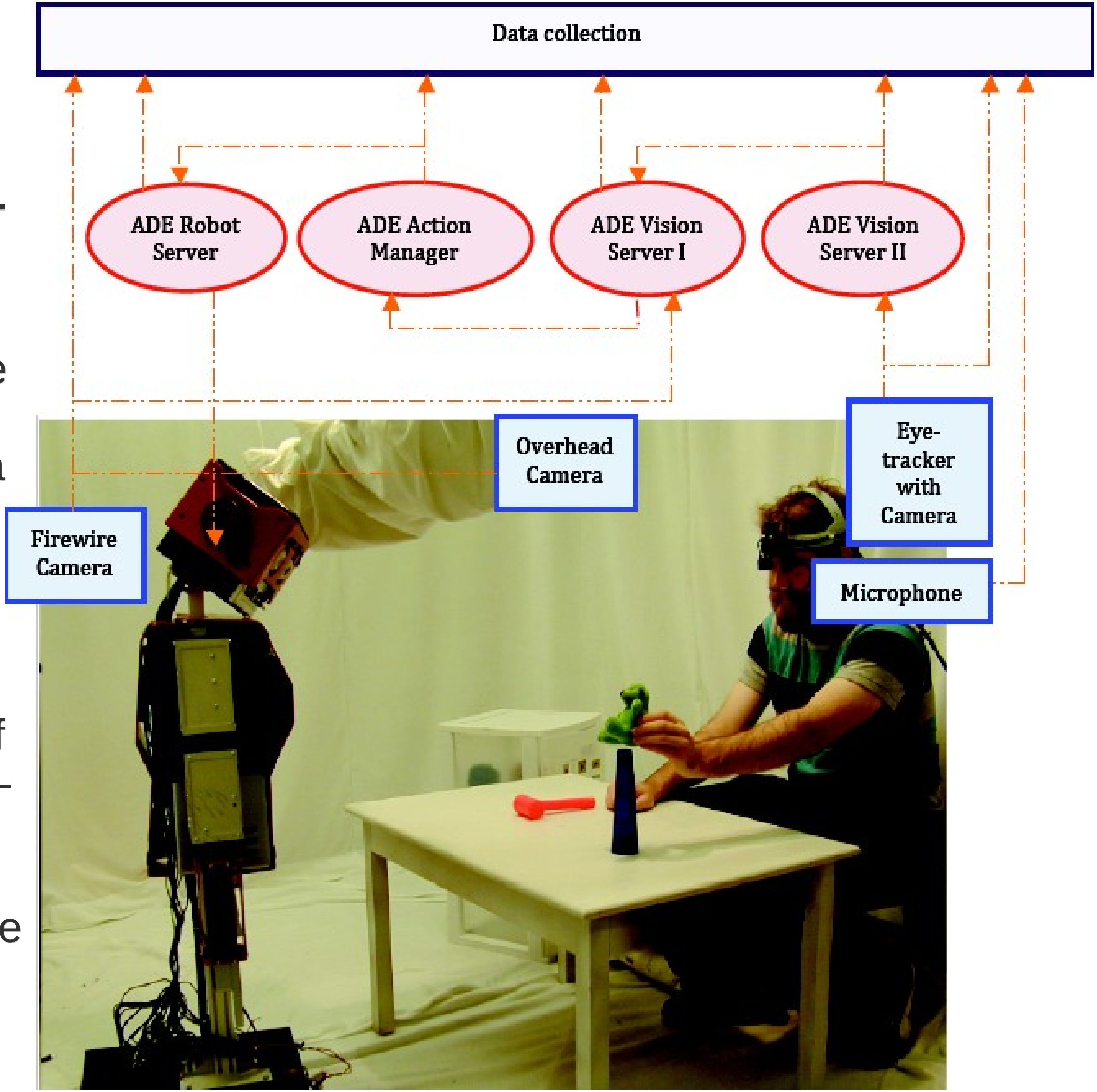
- Joint attention processes:
  - establishing and maintaining joint attention
  - breaking joint attention (through “abnormal attention”)
- Human attitudes about robots:
  - social facilitation and social inhibition (to probe agency)
  - robotic voice, social presence and gender differences
- Human reactions to autonomous robots in cooperative tasks:
  - human reactions to robot affect
  - human reactions to real vs simulated robots w/o autonomy
- Task-switching in human multi-tasking:
  - fNIRs-based adaption of robot autonomy
  - effects of real vs virtual robots on multi-tasking performance
- Philosophical and conceptual inquiry:
  - what it is like to be an agent/have a red experience?
  - the effects of “ethical robots” on human decision-making

Joint attention processes in parents and children

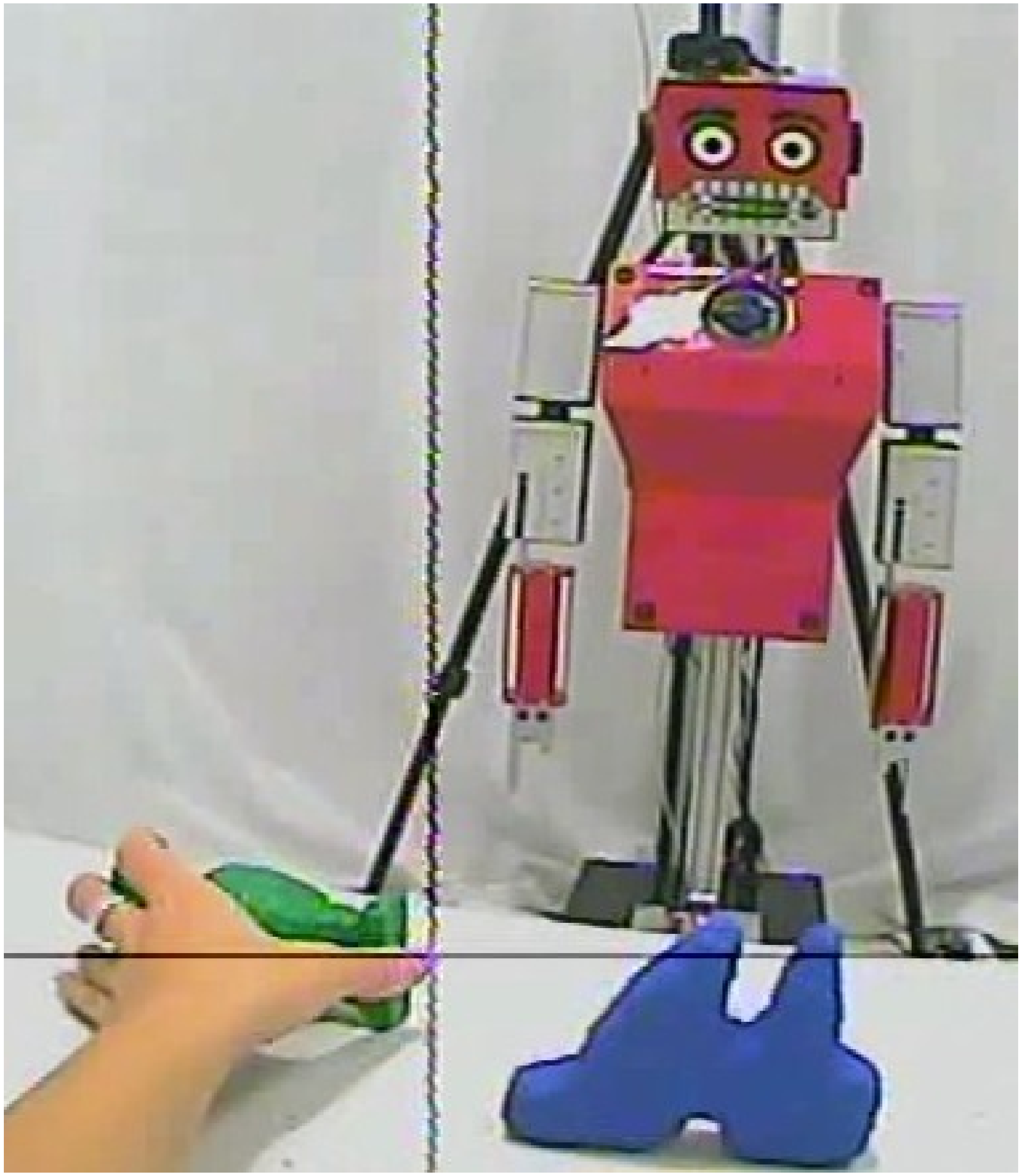
- Yu, Smith, Shen, Pereira, and Thomas (2009) studied the different dynamic structures of children's and parents' views of the events in the shared task of toy play and word learning
- Multi-modal data recording to obtain detailed time-course information



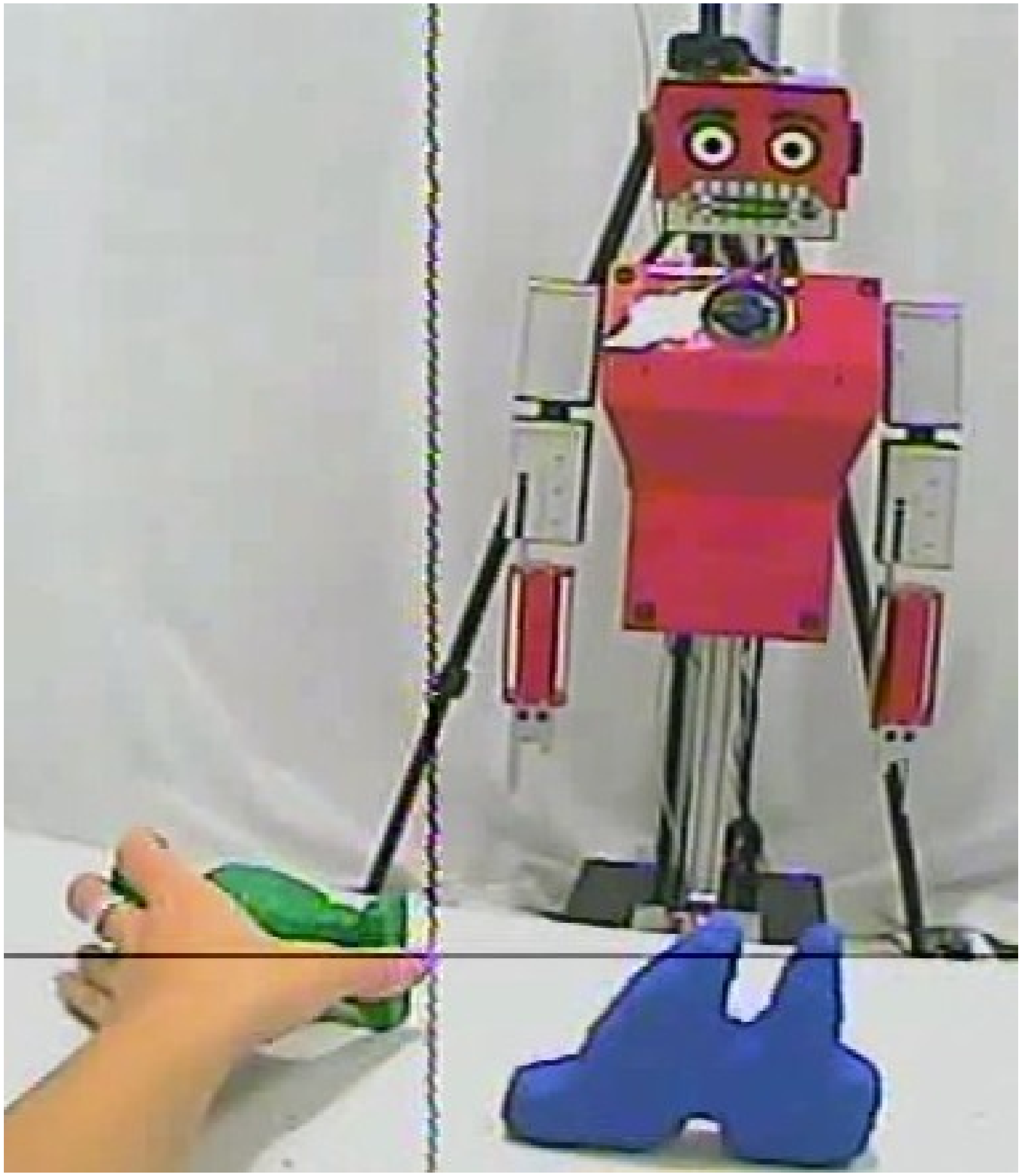
- Yu et al. (2010, 2011) replicated the experimental setup using a robot instead of human participant
- Required processing of real-time eye-gaze data and real-time reaction to the data (e.g., head moves)

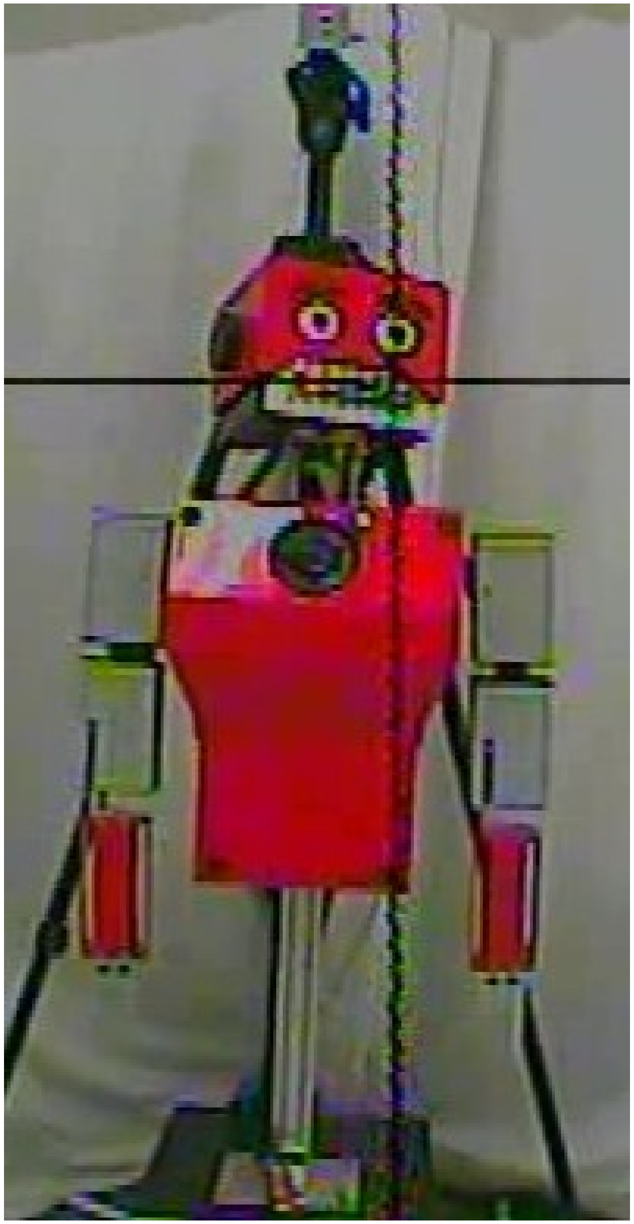


The "follow" condition

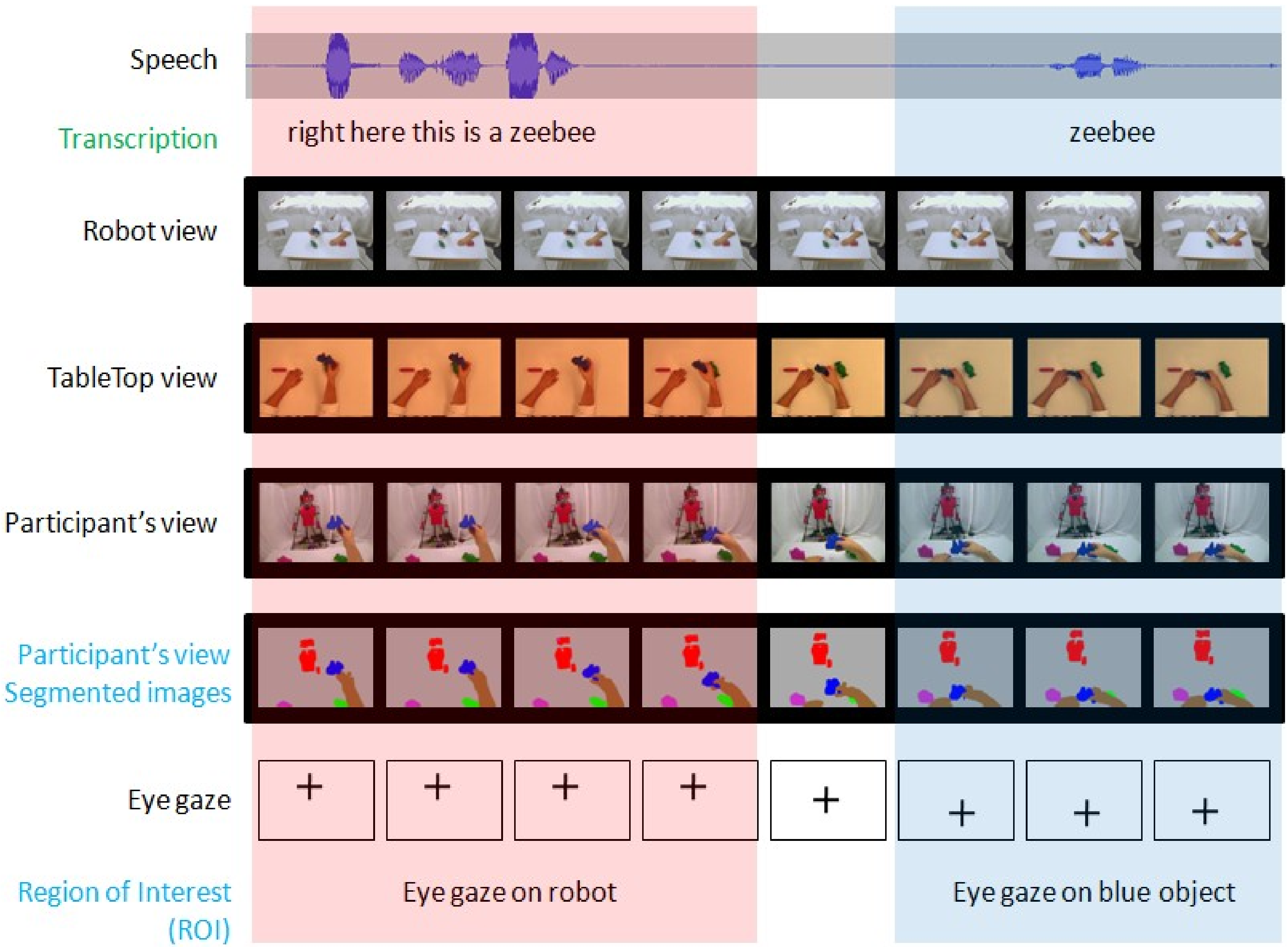


The “random” condition





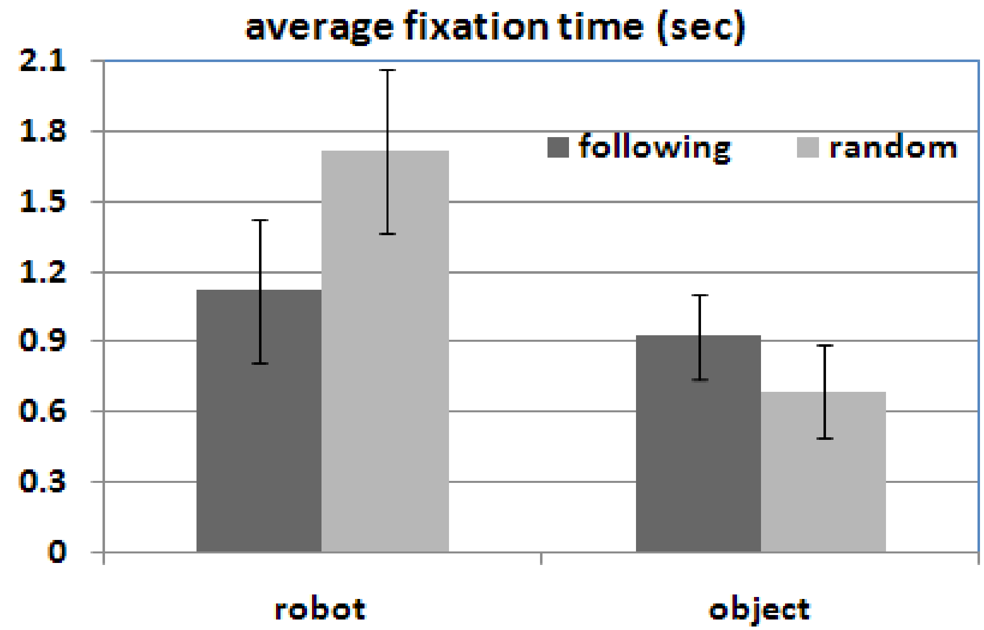
Multi-modal data collection



Eye fixations results

	following	random
number of attention switches (eye fixations)	53.61	55.8
average fixation length in seconds	0.96	1.16
number of robot looking fixations	22.32	21.75
average length of robot fixations in seconds (**)	1.11	1.72
longest fixation in seconds (**)	3.66	5.92

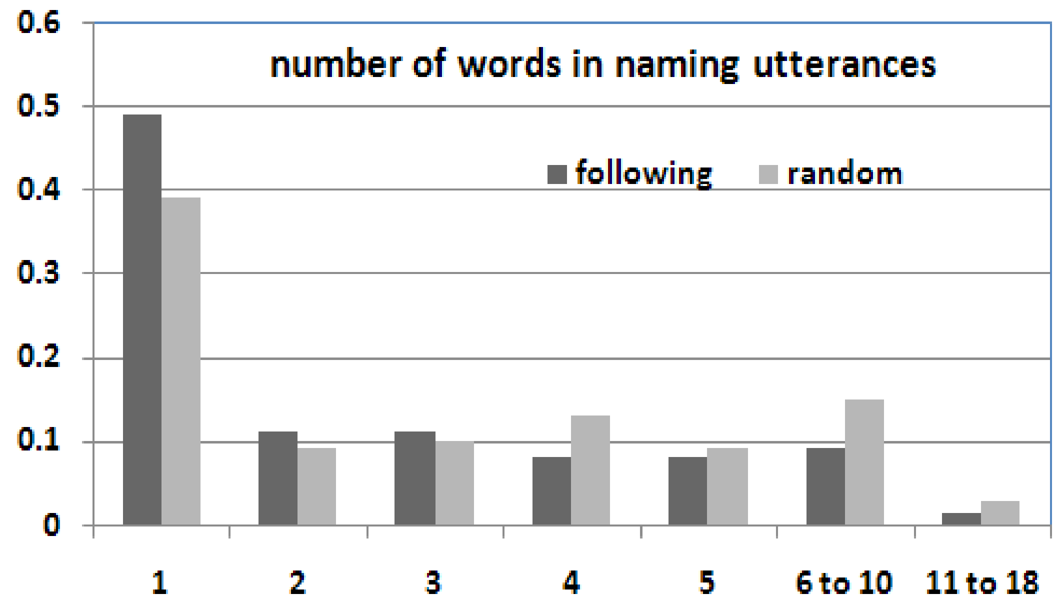
- Number of robot looking fixations is the same in both condition, so is the number of attention switches and robot fixations
- Participants in the random condition visually attended to the robot significantly longer (through longer eye fixations) than to objects and also longer than those in the following group



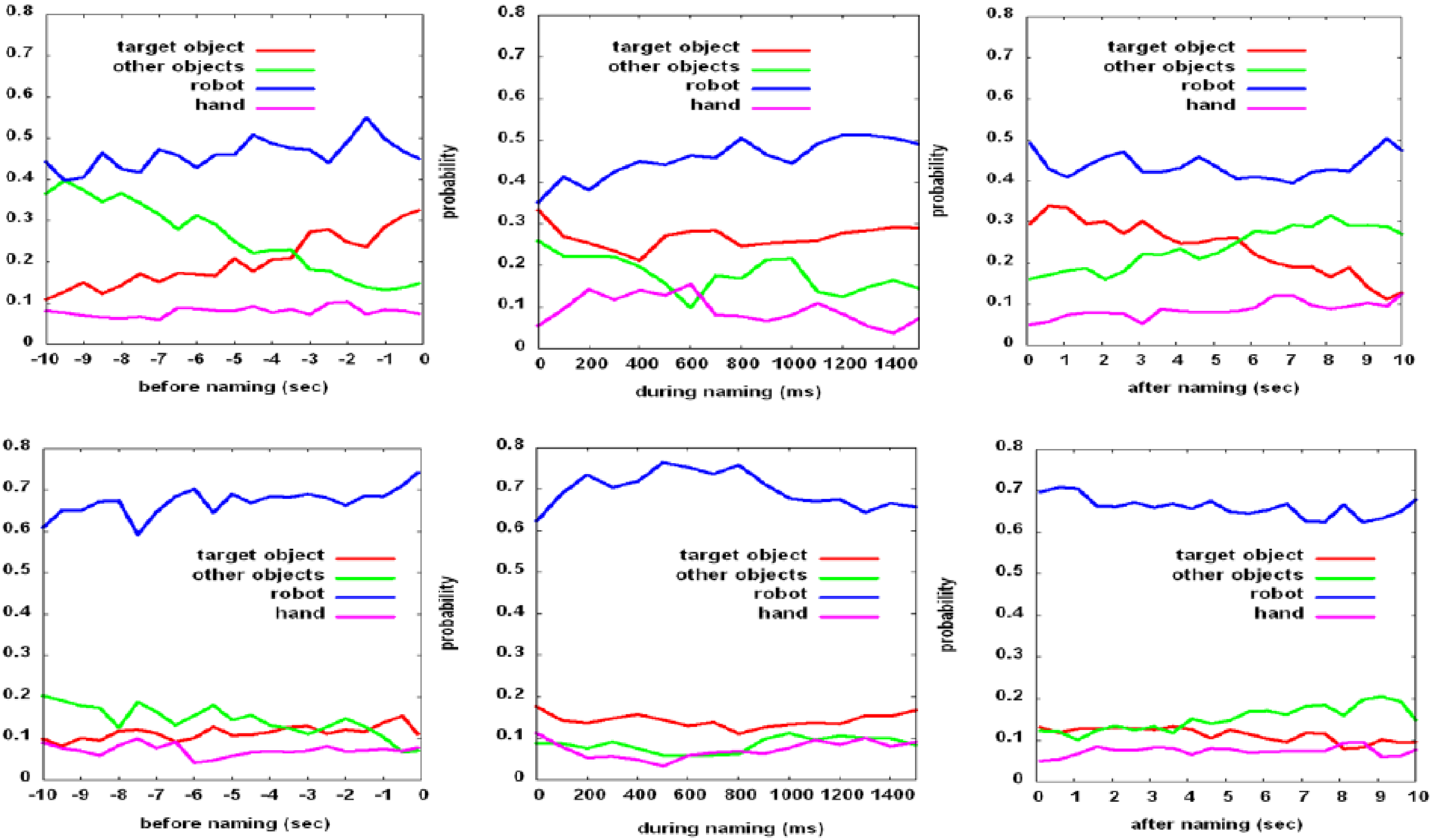
Naming utterances

	following	random
number of words	88	86
number of tokens (*)	394	459
number of utterances	114	121
words per utterance (*)	3.31	3.79
number of naming utterances (**)	48	60
proportion of naming utterances (**)	0.42	0.50

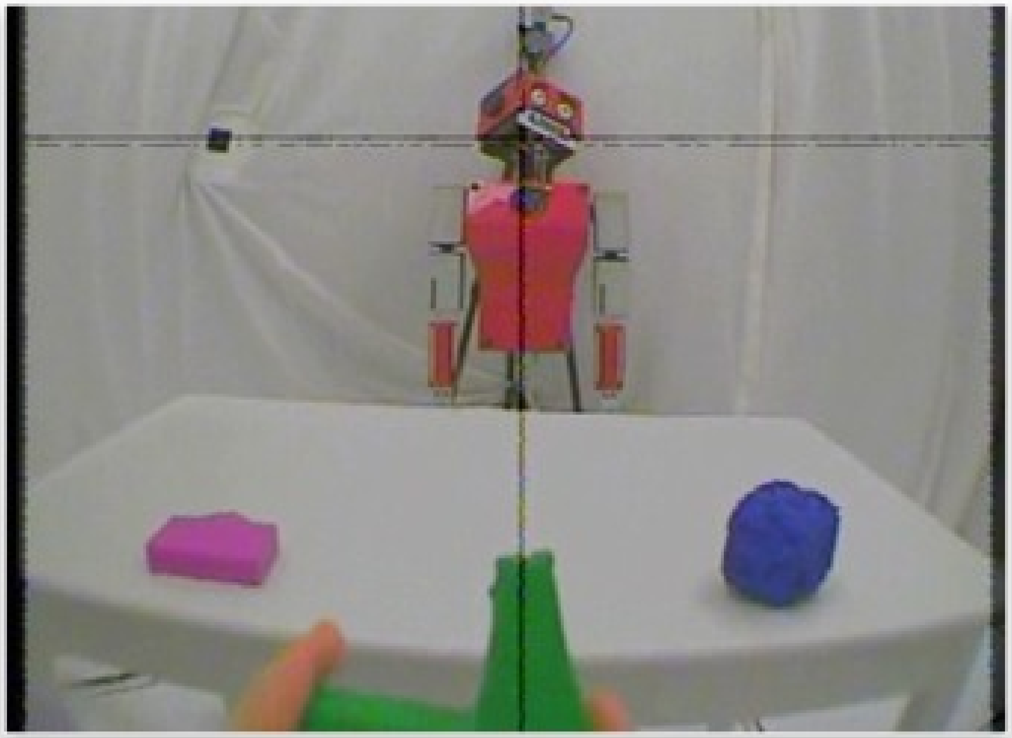
- Participants in the random condition produced significantly more naming utterances (containing object names) than those in the following group (60 versus 48); and this is not simply because the overall number of utterances differs



Temporal dynamics before, during and after naming events

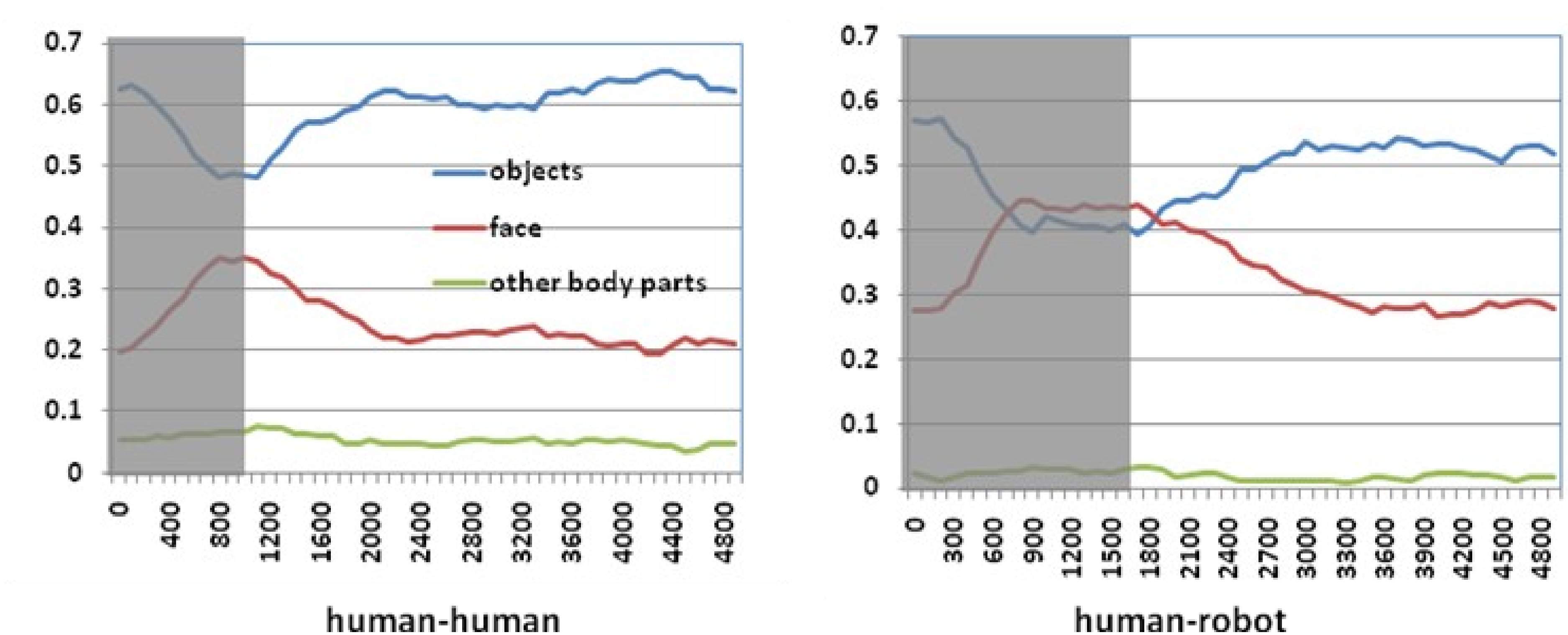


The “human robot” condition



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

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

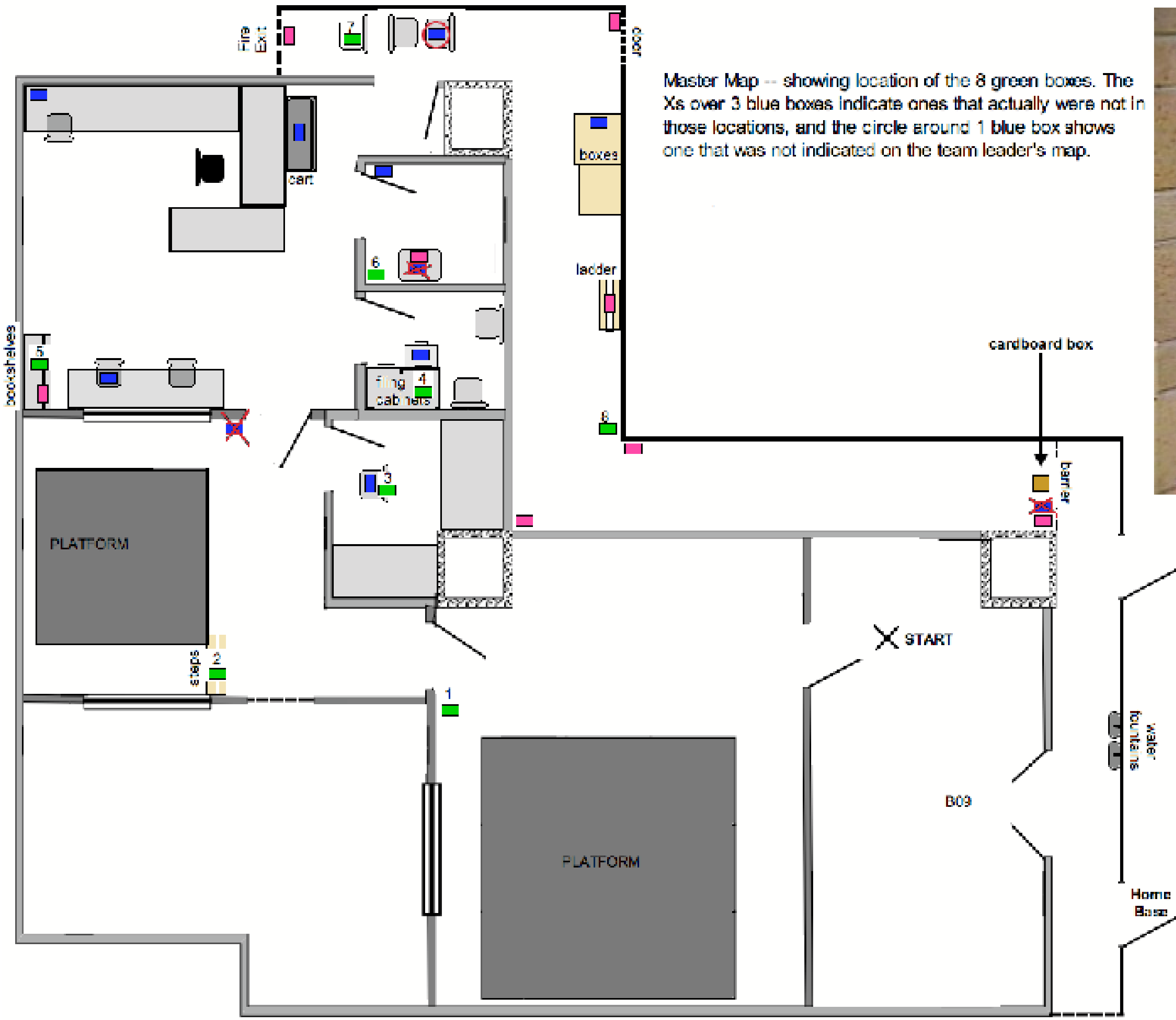
Robots as models: examples from our work

- Spoken natural language and dialogue interactions:
  - instructing and tasking in natural language
  - dialogue-based joint human-robot activities
- Introspection and self-awareness:
  - detecting faults and failures
  - detecting capabilities and possibilities
- Planning, reasoning, and problem solving in open worlds:
  - planning and reasoning with incopmlete knowledge
  - determining optimal policies in open worlds
- Knowledge-based learning:
  - one-shot learning new actions
  - one-shot learning of new plan operators
- Mental models, simulation, and counterfactual reasoning:
  - adverbial cues for inferring false beliefs
  - automatic inference from mental models

Robots as models: examples from our work

- Spoken natural language and dialogue interactions:
  - instructing and tasking in natural language
  - dialogue-based joint human-robot activities
- Introspection and self-awareness:
  - detecting faults and failures
  - detecting capabilities and possibilities
- Planning, reasoning, and problem solving in open worlds:
  - planning and reasoning with incomplete knowledge
  - determining optimal policies in open worlds
- Knowledge-based learning:
  - one-shot learning new actions
  - one-shot learning of new plan operators
- Mental models, simulation, and counterfactual reasoning:
  - adverbial cues for inferring false beliefs
  - automatic inference from mental models

The CReST task and corpus



Master Map -- showing location of the 8 green boxes. The Xs over 3 blue boxes indicate ones that actually were not in those locations, and the circle around 1 blue box shows one that was not indicated on the team leader's map.

Task-based HR dialogues

Robot has been going down one corridor and has stopped outside of a doorway.

H: Is there a hallway?
R: I see a hallway.
H: Okay, go down there.
R: Okay.

Question requires perceptual action, followed by human instruction with deictic reference which is carried out

The robot drives down the hallway. As it is driving down it notices a doorway, which it reports to the team leader, also acknowledging its position.

R: Okay, I'm now in the hallway. There is a doorway on the left.
H: Good, go through that doorway.
R: Okay.

The robot enters the room through the doorway and notices several yellow blocks. Since these are task-relevant, it reports them to the team leader.

R: I'm now inside the room. There are yellow blocks in boxes.
H: Get a yellow block from a blue box.
R: Okay.

The robot verifies that there is a yellow block in a blue box, approaches the blue box, and gets the yellow block.

Task-based HR dialogues

Robot has been going down one corridor and has stopped outside of a doorway.

H: Is there a hallway?
R: I see a hallway.
H: Okay, go down there.
R: Okay.

Updates to the team leader; ongoing active perceptual monitoring generates additional updates on task-relevant landmarks

The robot drives down the hallway. As it is driving down it notices a doorway, which it reports to the team leader, also acknowledging its position.

R: Okay, I'm now in the hallway. There is a doorway on the left.
H: Good, go through that doorway.
R: Okay.

The robot enters the room through the doorway and notices several yellow blocks. Since these are task-relevant, it reports them to the team leader.

R: I'm now inside the room. There are yellow blocks in boxes.
H: Get a yellow block from a blue box.
R: Okay.

The robot verifies that there is a yellow block in a blue box, approaches the blue box, and gets the yellow block.

Task-based HR dialogues

Robot has been going down one corridor and has stopped outside of a doorway.

H: Is there a hallway?
R: I see a hallway.
H: Okay, go down there.
R: Okay.

More updates after completed actions and results from task-based attentional bias; spatial reference resolution and action generation

The robot drives down the hallway. As it is driving down it notices a doorway, which it reports to the team leader, also acknowledging its position.

R: Okay, I'm now in the hallway. There is a doorway on the left.
H: Good, go through that doorway.
R: Okay.

The robot enters the room through the doorway and notices several yellow blocks. Since these are task-relevant, it reports them to the team leader.

R: I'm now inside the room. There are yellow blocks in boxes.
H: Get a yellow block from a blue box.
R: Okay.

The robot verifies that there is a yellow block in a blue box, approaches the blue box, and gets the yellow block.

Task-based HR dialogues

Robots as models: examples from our work

- Spoken natural language and dialogue interactions:
  - instructing and tasking in natural language
  - dialogue-based joint human-robot activities
- Introspection and self-awareness:
  - detecting faults and failures
  - detecting capabilities and possibilities
- Planning, reasoning, and problem solving in open worlds:
  - planning and reasoning with incomplete knowledge
  - determining optimal policies in open worlds
- Knowledge-based learning:
  - one-shot learning new actions
  - one-shot learning of new plan operators
- Mental models, simulation, and counterfactual reasoning:
  - adverbial cues for inferring false beliefs
  - automatic inference from mental models

Monitoring, introspecting, and discovering new capabilities

A robot is performing mission, when it gets a new order:

H: I have a new order for you. Go to Nav Point 4 and inspect all suspicious objects.

The robot immediately attempts to simulate the goal and determines that it cannot complete it because the “examine objects” perception action fails (due to the fact that the robot does not have a camera and therefore no vision processing component instantiated). It reports the problem to the human operator:

Tasking in natural language

R: I cannot achieve the goals because I do not have a way to examine objects.
H: Okay, come back and I will get you a camera.
R: Okay.

The robot postpones the goals from the order and returns to the base where human operator installs camera. The camera causes a vision processing component to be started automatically, which, in turn, allows the robot to notice that it can now examine objects.

R: I am now able to examine objects, resuming postponed goals.

The robot then moves goes to Nav Point 4 and starts looking for suspicious objects. It detects a suspicious unattended object, reports it to the operator and starts to inspect it.

R: I found an unattended crate, moving to investigate.

Monitoring, introspecting, and discovering new capabilities

A robot is performing mission, when it gets a new order:

H: I have a new order for you. Go to Nav Point 4 and inspect all suspicious objects.

Simulation of task performance uncovers that the task cannot be performed with the robot's current configuration

The robot immediately attempts to simulate the goal and determines that it cannot complete it because the "examine objects" perception action fails (due to the fact that the robot does not have a camera and therefore no vision processing component instantiated). It reports the problem to the human operator:

R: I cannot achieve the goals because I do not have a way to examine objects.
H: Okay, come back and I will get you a camera.
R: Okay.

The robot postpones the goals from the order and returns to the base where human operator installs camera. The camera causes a vision processing component to be started automatically, which, in turn, allows the robot to notice that it can now examine objects.

R: I am now able to examine objects, resuming postponed goals.

The robot then moves goes to Nav Point 4 and starts looking for suspicious objects. It detects a suspicious unattended object, reports it to the operator and starts to inspect it.

R: I found an unattended crate, moving to investigate.

Monitoring, introspecting, and discovering new capabilities

A robot is performing mission, when it gets a new order:

H: I have a new order for you. Go to Nav Point 4 and inspect all suspicious objects.

The robot immediately attempts to simulate the goal and determines that it cannot complete it because the "examine objects" perception action fails (due to the fact that the robot does not have a camera and therefore no vision processing component instantiated). It reports the problem to the human operator:

R: I cannot achieve the goals because I do not have a way to examine objects.
H: Okay, come back and I will get you a camera.
R: Okay.

The robot postpones the goals from the order and returns to the base where human operator installs camera. The camera causes a vision processing component to be started automatically, which, in turn, allows the robot to notice that it can now examine objects.

Automatic detection of new sensory device and auto-configuration of processing components, reflection on functional capabilities

R: I am now able to examine objects, resuming postponed goals.

The robot then moves goes to Nav Point 4 and starts looking for suspicious objects. It detects a suspicious unattended object, reports it to the operator and starts to inspect it.

R: I found an unattended crate, moving to investigate.

Monitoring, introspecting, and discovering new capabilities

A robot is performing mission, when it gets a new order:

H: I have a new order for you. Go to Nav Point 4 and inspect all suspicious objects.

The robot immediately attempts to simulate the goal and determines that it cannot complete it because the “examine objects” perception action fails (due to the fact that the robot does not have a camera and therefore no vision processing component instantiated). It reports the problem to the human operator:

R: I cannot achieve the goals because I do not have a way to examine objects.
H: Okay, come back and I will get you a camera.
R: Okay.

The robot postpones the goals from the order and returns to the base where human operator installs camera. The camera causes a vision processing component to be started automatically, which, in turn, allows the robot to notice that it can now examine objects.

R: I am now able to examine objects, resuming postponed goals.

The robot then moves goes to Nav Point 4 and starts looking for suspicious objects. It detects a suspicious unattended object, reports it to the operator and starts to inspect it.

R: I found an unattended crate, moving to investigate.

Continuation of previously postponed goal, target dection and automatic notification of operator

Monitoring, introspecting, and discovering new capabilities

Robots as models: examples from our work

- Spoken natural language and dialogue interactions:
  - instructing and tasking in natural language
  - dialogue-based joint human-robot activities
- Introspection and self-awareness:
  - detecting faults and failures
  - detecting capabilities and possibilities
- Planning, reasoning, and problem solving in open worlds:
  - planning and reasoning with incomplete knowledge
  - determining optimal policies in open worlds
- Knowledge-based learning:
  - one-shot learning new actions
  - one-shot learning of new plan operators
- Mental models, simulation, and counterfactual reasoning:
  - adverbial cues for inferring false beliefs
  - automatic inference from mental models

Tasking through NL dialogues, introspection and dynamic autonomy

- *Backdrop:* autonomous urban search and rescue missing
- *Goals:* **hard** (e.g., follow team leader's commands) and **soft** (e.g., search for wounded people and report their location)
- Robot has **prior knowledge** that wounded people might be hidden in rooms
- Robot **counterfactually** determines that for it to be able to find people in rooms, it must first find doorways to rooms
- Robot **schedules** perception actions for doorways while it is executing other commands
- Robot **attempts to satisfy soft goals** if goal satisfaction does not violate hard goals and if the expected utility of soft goals >0
- Robot **monitors goals for consistency** and communicates with human team leader about any potential inconsistencies

Tasking through NL dialogues, introspection and dynamic autonomy

H: What are your orders?
R: My orders are to listen for instructions.
H: OK, here are your orders. Go to the end of the hallway.
R: OK.
H: Um, keep the lights off at all times.
R: OK.
H: And try to report the locations of wounded people.
R: OK.
H: Proceed.
R: OK
(Robot starts down the hallway, sees first door, enters room, scans room.)
R: There is a wounded person in the first room on the left.
(Robot leaves room, continues down the hallway. IR camera stops working).
R: I have lost contact with the infrared camera. Switching to the regular camera.
(Robot continues down the hallway, sees second door, enters room, tries to scan.)
R: I need to activate a light, but I have a goal to keep lights off. What shall I do?
H: Um, cancel keep lights off.
R: OK, goal cancelled.
(Robot turns light on, scans room. Robot leaves room, continues down the hallway.)
H: What are your orders?
R: My orders are to listen for instructions, go to the end of the hallway and try
to report the locations of wounded people.
H: You have one minute left to get to the end of the hallway.
R: OK.
(Robot skips remaining rooms.)

Tasking through NL dialogues, introspection and dynamic autonomy

H: What are your orders?
R: My orders are to listen for instructions

Query robot goal status

H: OK, here are your orders. Go to the end of the hallway.
R: OK.
H: Um, keep the lights off at all times.
R: OK.
H: And try to report the locations of wounded people.
R: OK.
H: Proceed.
R: OK
(Robot starts down the hallway, sees first door, enters room, scans room.)
R: There is a wounded person in the first room on the left.
(Robot leaves room, continues down the hallway. IR camera stops working).
R: I have lost contact with the infrared camera. Switching to the regular camera.
(Robot continues down the hallway, sees second door, enters room, tries to scan.)
R: I need to activate a light, but I have a goal to keep lights off. What shall I do?
H: Um, cancel keep lights off.
R: OK, goal cancelled.
(Robot turns light on, scans room. Robot leaves room, continues down the hallway.)
H: What are your orders?
R: My orders are to listen for instructions, go to the end of the hallway and try to report the locations of wounded people.
H: You have one minute left to get to the end of the hallway.
R: OK.
(Robot skips remaining rooms.)

Tasking through NL dialogues, introspection and dynamic autonomy

H: What are your orders?
R: My orders are to listen for instructions.
H: OK, here are your orders. Go to the end of the hallway.
R: OK.
H: Um, keep the lights off at all times.
R: OK.
H: And try to report the locations of wounded people.
R: OK.
H: Proceed.
R: OK

New hard and soft goals for the robot in natural language

(Robot starts down the hallway, sees first door, enters room, scans room.)
R: There is a wounded person in the first room on the left.
(Robot leaves room, continues down the hallway. IR camera stops working).
R: I have lost contact with the infrared camera. Switching to the regular camera.
(Robot continues down the hallway, sees second door, enters room, tries to scan.)
R: I need to activate a light, but I have a goal to keep lights off. What shall I do?
H: Um, cancel keep lights off.
R: OK, goal cancelled.
(Robot turns light on, scans room. Robot leaves room, continues down the hallway.)
H: What are your orders?
R: My orders are to listen for instructions, go to the end of the hallway and try to report the locations of wounded people.
H: You have one minute left to get to the end of the hallway.
R: OK.
(Robot skips remaining rooms.)

Tasking through NL dialogues, introspection and dynamic autonomy

Robot starts goal execution, determines opportunity to meet a soft goal, pursues it, and generates an NL report after discovery of a target

H: What are your orders?
R: My orders are to listen for instructions.
H: OK, here are your orders. Go to the end of the hallway.
R: OK.
H: Um, keep the lights off at all times.
R: OK.
H: And try to report the locations of wounded people.
R: OK.
H: Proceed.
R: OK.
(Robot starts down the hallway, sees first door, enters room, scans room.)
R: There is a wounded person in the first room on the left.
(Robot leaves room, continues down the hallway. IR camera stops working).
R: I have lost contact with the infrared camera. Switching to the regular camera.
(Robot continues down the hallway, sees second door, enters room, tries to scan.)
R: I need to activate a light, but I have a goal to keep lights off. What shall I do?
H: Um, cancel keep lights off.
R: OK, goal cancelled.
(Robot turns light on, scans room. Robot leaves room, continues down the hallway.)
H: What are your orders?
R: My orders are to listen for instructions, go to the end of the hallway and try
to report the locations of wounded people.
H: You have one minute left to get to the end of the hallway.
R: OK.
(Robot skips remaining rooms.)

Tasking through NL dialogues, introspection and dynamic autonomy

H: What are your orders?
R: My orders are to listen for instructions.
H: OK, here are your orders. Go to the end of the hallway.
R: OK.
H: Um, keep the lights off at all times.
R: OK.
H: And try to report the locations of wounded people.
R: OK.
H: Proceed.
R: OK
(Robot starts down the hallway, sees first door, enters room, scans room.)
R: There is a wounded person in the first room on the left.
(Robot leaves room, continues down the hallway. IR camera stops working).
R: I have lost contact with the infrared camera. Switching to the regular camera.
(Robot continues down the hallway, sees second door, enters room, tries to scan.)
R: I need to activate a light, but I have a goal to keep lights off. What shall I do?
H: Um, cancel keep lights off.
R: OK, goal cancelled.
(Robot turns light on, scans room. Robot leaves room, continues down the hallway.)
H: What are your orders?
R: My orders are to listen for instructions, go to the end of the hallway and try to report the locations of wounded people.
H: You have one minute left to get to the end of the hallway.
R: OK.
(Robot skips remaining rooms.)

Adjustment of system configuration based on internal monitoring

Tasking through NL dialogues, introspection and dynamic autonomy

H: What are your orders?
R: My orders are to listen for instructions.
H: OK, here are your orders. Go to the end of the hallway.
R: OK.
H: Um, keep the lights off at all times.
R: OK.
H: And try to report the locations of wounded people.
R: OK.
H: Proceed.
R: OK
(Robot starts down the hallway, sees first door, enters room, scans room.)
R: There is a wounded person in the first room on the left.
(Robot leaves room, continues down the hallway. IR camera stops working).
R: I have lost contact with the infrared camera. Switching to the regular camera.
(Robot continues down the hallway, sees second door, enters room, tries to scan.)
R: I need to activate a light, but I have a goal to keep lights off. What shall I do?
H: Um, cancel keep lights off.
R: OK, goal cancelled.
(Robot turns light on, scans room. Robot leaves room, continues down the hallway.)
H: What are your orders?
R: My orders are to listen for instructions, go to the end of the hallway and try
to report the locations of wounded people.
H: You have one minute left to get to the end of the hallway.
R: OK.
(Robot skips remaining rooms.)

Fault detection and action to remedy it, which is in conflict with goal – robot seeks permission from human to proceed, which requires a hard goal to be cancelled

Tasking through NL dialogues, introspection and dynamic autonomy

H: What are your orders?
R: My orders are to listen for instructions.
H: OK, here are your orders. Go to the end of the hallway.
R: OK.
H: Um, keep the lights off at all times.
R: OK.
H: And try to report the locations of wounded people.
R: OK.
H: Proceed.
R: OK
(Robot starts down the hallway, sees first door, enters room, scans room.)
R: There is a wounded person in the first room on the left.
(Robot leaves room, continues down the hallway. IR camera stops working).
R: I have lost contact with the infrared camera. Switching to the regular camera.
(Robot continues down the hallway, sees second door, enters room, tries to scan.)
R: I need to activate a light, but I have a goal to keep lights off. What shall I do?
H: Um, cancel keep lights off.
R: OK, goal cancelled.
(Robot turns light on, scans room. Robot leaves room, continues down the hallway.)

Query of goal status, update does not contain cancelled goal any longer

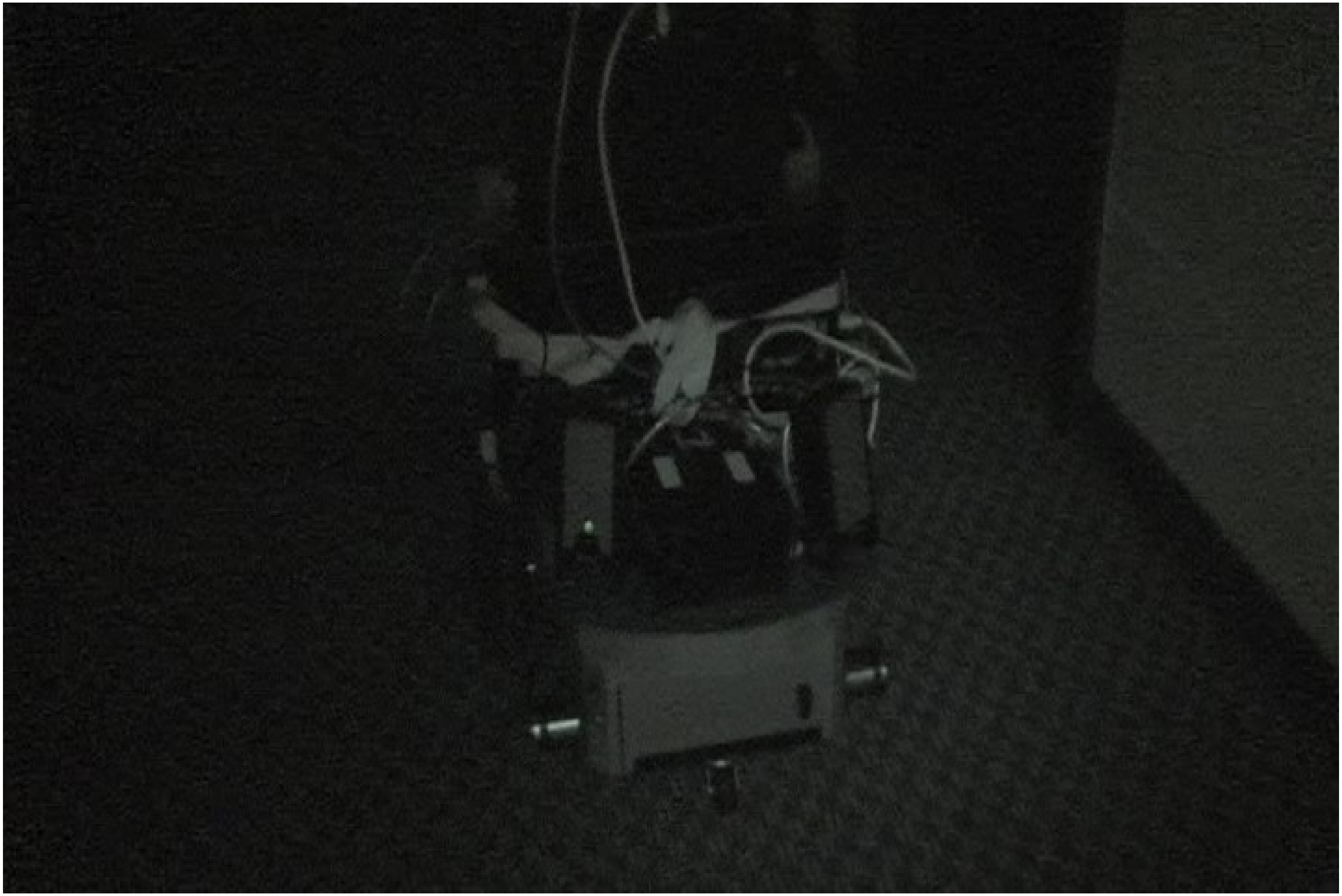
H: What are your orders?
R: My orders are to listen for instructions, go to the end of the hallway and try to report the locations of wounded people.
H: You have one minute left to get to the end of the hallway.
R: OK.
(Robot skips remaining rooms.)

Tasking through NL dialogues, introspection and dynamic autonomy

H: What are your orders?
R: My orders are to listen for instructions.
H: OK, here are your orders. Go to the end of the hallway.
R: OK.
H: Um, keep the lights off at all times.
R: OK.
H: And try to report the locations of wounded people.
R: OK.
H: Proceed.
R: OK
(Robot starts down the hallway, sees first door, enters room, scans room.)
R: There is a wounded person in the first room on the left.
(Robot leaves room, continues down the hallway. IR camera stops working.)
R: I have lost contact with the infrared camera. Switching to the regular camera.
(Robot continues down the hallway, sees second door, enters room, tries to scan.)
R: I need to activate a light, but I have a goal to keep lights off. What shall I do?
H: Um, cancel keep lights off.
R: OK, goal cancelled.
(Robot turns light on, scans room. Robot leaves room, continues down the hallway.)
H: What are your orders?
R: My orders are to listen for instructions, go to the end of the hallway and try to report the locations of wounded people.
H: You have one minute left to get to the end of the hallway.
R: OK.
(Robot skips remaining rooms.)

Changes in goal priorities, which trigger re-planning and subsequent skipping of soft goals to meet hard goal

Tasking through NL dialogues, introspection and dynamic autonomy



Conclusions and future work

- We argued that robots can serve a **dual role** in cognitive science, both as experimental tools and embodied models
- Then we introduced the DIARC architecture framework that has been used for performing HRI experiments and for implementing embodied models of situated interactions
- We presented different examples from our work (using DIARC) both for **collecting data about humans behaviors** and for **modeling human cognitive capabilities**
- Currently, we are working on an even tighter human-like integration between different functional components in the architecture (e.g., incremental vision and incremental natural language processing, incremental planning and reasoning) as well as novel biologically plausible components (e.g., for visual attentional biases and speech processing)

Acknowledgments and links

- Collaborators on the work presented:
  - Faculty: Kathleen Eberhard (Notre Dame), Sandra Kuebler and Chen Yu (Indiana U), Rao Kambhampati and Chitta Baral (ASU), Roni Khardon (Tufts U)
  - Postdocs and students: Paul Schermerhorn (IU/Tufts), Virgil Andronache (ND), Jim Kramer (ND), Rehj Cantrell (IU), Zach Haga (IU), Tim Brick (ND), Richard Veale (IU), Kyle Carter (IU), and other students of collaborators
- Funding provided in part by the **US National Science Foundation** and the **US Office of Naval Research**

- Publications available online at:

  http://hrilab.cs.tufts.edu/publications/
- Software available online at:

  http://hrilab.cs.tufts.edu/software/
  http://ade.sourceforge.net/