Human-Robot Interaction

Communication

Organisms communicate via a number of **modalities**: visual, auditory, haptic (tactile), olfactory, and gustatory.

Modalities can include a number of **channels**, for example, visual signals can be displayed in the face, body, and hands.

According to many authors like (Hogan and Stubbs, 2003), more than sixty percent of the communication between two people or between one speaker and a group of listeners is nonverbal.

Communication

Communication is an interactive process where messages are continuously exchanged between sender and receiver.

Any definition of communication in HRI needs to account for non-human (and potentially non-humanlike) message transmission and communication behavior that is intentional or unintentional.

Communication is an interactive process whereby a receiver (i.e., a person) assigns meaning to one or more stimuli that are transmitted intentionally or unintentionally by a sender (i.e., a robot).

Non-verbal Interaction

Nonverbal communication in HRI as any type of communication between people and robots that does not involve words but includes nonverbal utterances and behaviors.

Nonverbal communication, which is generally defined as any transfer of messages that does not involve the use of words. This also includes sounds that are not words, such as back channeling ("uh" and "ah") as well as the pitch, tone, and intonation of the voice.

People constantly and seemingly automatically pick up on a variety of nonverbal cues while interacting. These cues are used to interpret the nuances of meaning, emotion, and intention in others.

https://www.youtube.com/watch?v=pZCSp2FSKDE&list=PPSV

Non-verbal Interaction

Ekman and Friesen (1969) have suggested that there are five ways in which nonverbal behavior qualifies the meaning of the verbal content:

(a) repetition,

(b) contradiction,

(c) complementary function (e.g., praise may be accompanied by a smile or anger with a clenched fist),

(d) accentuation, and

(e) regulation (e.g., eye contact influences conversational turn talking).

Functions of Nonverbal Cues

- Nonverbal cues allow people to communicate important information “between the lines.”
- Through nonverbal communication, people can signal mutual understanding, shared goals, and common ground.
- They can communicate thoughts, emotions, and attention. And they can do so in a more subtle, indirect manner than through verbal expression.
- Many of the nonverbal signals we convey are expressed automatically without much thought or are even entirely unconsciously. Therefore, nonverbal cues are often believed to be unaltered and more genuine, revealing people's “true” attitudes.

See: https://socialsci.libretexts.org/Bookshelves/Communication/Interpersonal_Communication/I.C.A.T_Interpersonal_Communication_Abridged_Textbook_(Gerber_and_Murphy)/06%3A_Nonverbal_Communication/6.02%3A_Functions_of_Nonverbal_Communication

Functions of Nonverbal Cues

So How does it affect HRI?

Functions of Nonverbal Cues

So How does it affect HRI?

- Nonverbal cues *produced by people* when interacting with a robot can indicate whether a person is enjoying the interaction and whether the person likes the robot or not.
- HRI may also be affected by *the way robots produce* nonverbal cues.
    - A robot that displays nonverbal behavior (such as turning its head towards human during conversation) will make the interaction seem more natural and smooth. When it is missing we can sense that something is going wrong, even though it might be difficult to pinpoint exactly what is missing.
    - Interaction can appear awkward when the robot produces gestures that do not match the rhythm or meaning of its speech or when it does not respond appropriately to people's nonverbal cues.

Functions of Nonverbal Cues

Social cue is typically used to describe a set of temporal changes in neuromuscular and physiological activity that last for short intervals of time (milliseconds to minutes).

Human ***social signals*** are acts or structures that influence the behavior or internal state of other individuals, that evolve because of that effect, and that are effective because the perceivers response has also evolved; signals may or may not convey conceptual information or meaning, [Mehu and Scherer (2012)]. Examples of social signals are: attention, empathy, politeness, flirting and agreement.

Functions of Nonverbal Cues

To recognize an specific social signal we need to detect and analyze multiple social cues which occurs in our brains regularly and most of time unconsciously.

	Social Signals						
Social Cues	emotion	personality	status	dominance	persuasion	regulation	rapport
Physical appearance							
height			✓	✓			
attractiveness		✓	✓	✓	✓		✓
body shape		✓		✓			
Gesture and posture							
hand gestures	✓				✓	✓	✓
posture	✓	✓	✓	✓	✓	✓	✓
walking		✓	✓	✓			
Face and eyes behaviour							
facial expressions	✓	✓	✓	✓	✓	✓	✓
gaze behaviour	✓	✓	✓	✓	✓	✓	✓
focus of attention	✓	✓			✓	✓	✓
Vocal behaviour							
prosody	✓	✓			✓		
turn taking			✓	✓		✓	✓
vocalizations	✓	✓		✓	✓	✓	✓
silence							✓
Space and Environment							
distance		✓	✓		✓		✓
seating arrangement				✓			✓

Types of nonverbal interaction

Although we exhibit and experience nonverbal cues in several modalities at once, such as sound, movement, and gaze, it might be worthwhile to consider each channel of communication separately when trying to implement nonverbal signals into HRI. Common nonverbal cues are as follows

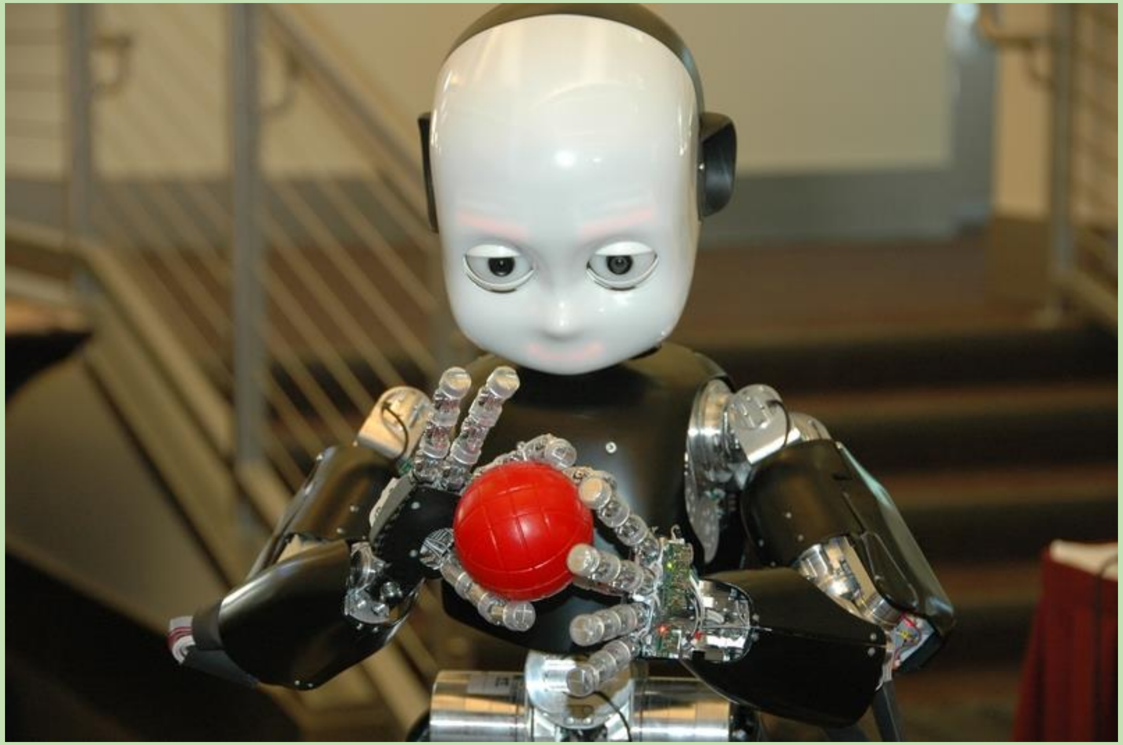
1. Gaze and eye movement
2. Gesture
3. Mimicry and Imitation
4. Touch
5. Posture and movement
6. Interaction rhythm and timing

Gaze and Eye movement

Eyes are a window to the soul, or in this case, they unconsciously reveal how much you like your interaction partner.

Pupil dilation is controlled by the autonomic nervous system, as are uncontrollable reactions such as an increase in heart rate or goose bumps.

The eyes of robots are often designed to pitch and yaw, allowing a robot to use gaze as an effective communication channel.



https://www.youtube.com/watch?v=-bBHT158E0s&t=136s

Gaze and Eye movement

Gaze signals some *social functions* such as

- interest,
- understanding,
- Attention during a conversation,
- turn-taking* in interactions and
- people's ability and willingness to follow the conversation.

* Turn-taking occurs in a conversation when one person listens while the other person speaks. As a conversation progresses, the listener and speaker roles are exchanged back and forth (a circle of discussion).

Gaze and Eye movement

Gaze signals some *social functions* such as

- interest,
- understanding,
- Attention during a conversation,
- turn-taking in interactions and
- people's ability and willingness to follow the conversation.

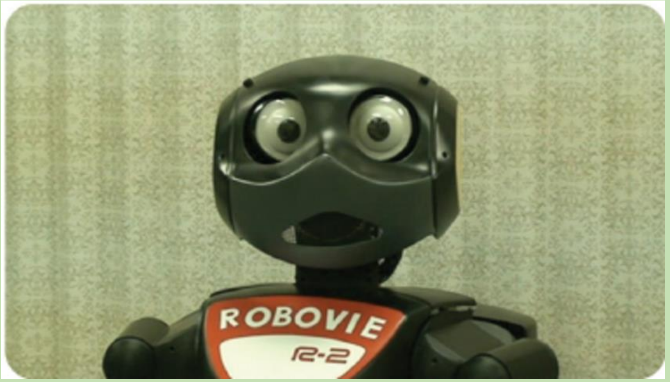
Gaze and eye movement facilitates some *functional interaction* such as

- handing an object to someone
- calling someone's attention to the next tool needed in a task.

Footing In Human-Robot Conversations: How Robots Might Shape Participant Roles Using Gaze Cues. Mutlu et al. (2009)

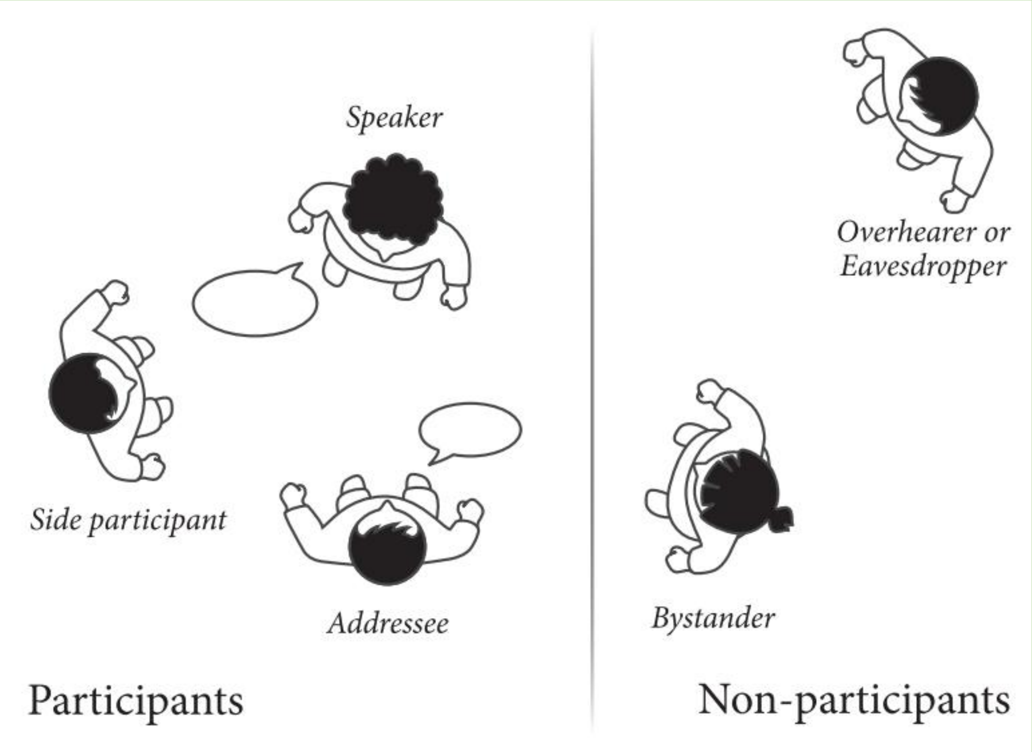
- During conversations, speakers establish their and others' participant roles (who participates in the conversation and in what capacity)—or "footing" as termed by Goffman—using gaze cues.

In this paper, they studied how a robot can establish the participant roles of its conversational partners using these cues. They designed a set of gaze behaviors for Robovie to signal three kinds of participant roles: addressee, bystander, and over hearer. They evaluated their design in a controlled laboratory experiment with 72 subjects in 36 trials. In three conditions, the robot signaled to two subjects, only by means of gaze, the roles of (1) two addressees, (2) an addressee and a bystander, or (3) an addressee and an over hearer.



Conversational Gaze Mechanisms for Humanlike Robots. Mutlu et al. (2012)

They developed models of three conversational gaze mechanisms; role-signaling, turn-taking, and topic signaling; integrated these models into a conversational system and implemented this system into a humanlike robot. The evaluation of the designed system assessed the effectiveness of these models in signaling three participant roles, addressee, bystander, and over hearer, and the social and cognitive effects of these participant roles on subjects.

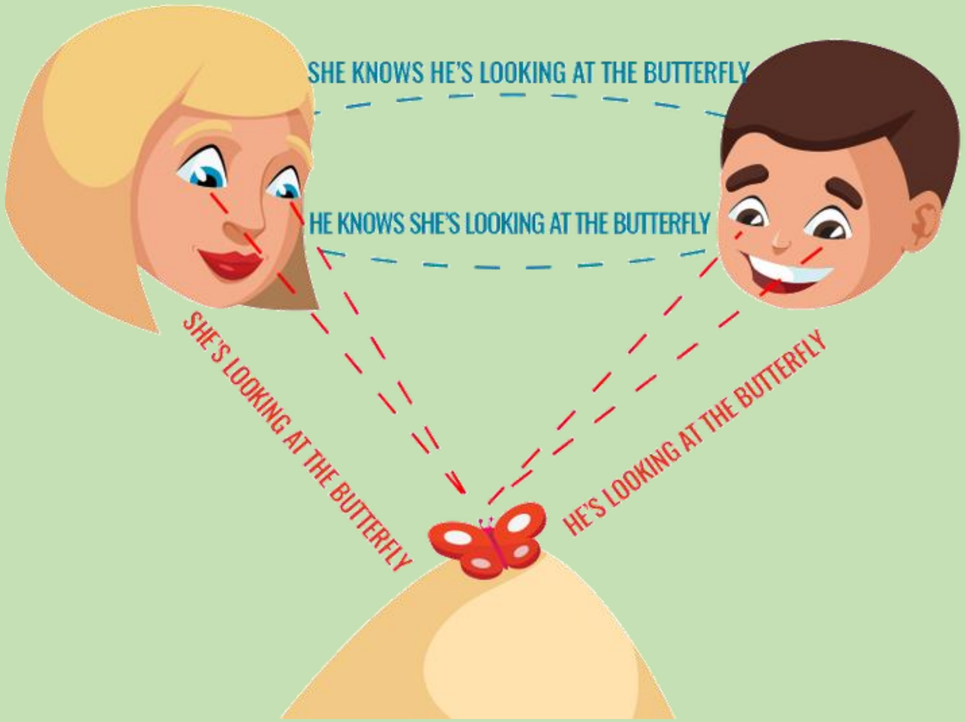


In a controlled laboratory environment with 72 subjects in 36 trials, they found that subjects correctly interpreted 99% of the turn-yielding signals and took 97% of the turns that the robot yielded to them.

Gaze and Eye movement

A particularly well-established component of gaze behavior in human interaction is joint attention.

Joint attention (aka shared attention) refers to interaction partners attending to the same area or object at the same time.



Gaze and Eye movement

Significances:

- The ability to attend to the same object at the same time with an adult caregiver is an important prerequisite for infants' ability to learn new words and behaviors (Yu and Smith, 2013), whereas the inability to perform joint attention can lead to developmental difficulties (Charman et al., 2000).
- Joint attention in adult communication can also signify interest and deep involvement in the interaction and is important for collaborative tasks where actors need to coordinate their activities.
- To achieve joint attention, the timing and synchrony of gaze behavior are important aspects to consider.

Gesture

Following speech, gesturing is perhaps the most apparent way of providing information during an interaction. Gesture can often be categorized based on their role in communication.

Classification of gesture:

Deictic gestures refer to pointing to specific things in the environment and can be important for establishing joint attention.

Iconic gestures often go along with speech, further supporting and illustrating what is being said. For example, opening your arms wide while saying you are holding a big ball would be an iconic gesture, as would smoothly moving your hand upward while explaining how your airplane took off.

Gesture

Classification of gesture:

Symbolic gestures, such as waving for hello or goodbye, can carry their own meaning, with or without accompanying speech.

Finally, beat gestures are used to go along with the rhythm of speech and look like moving one's arms while speaking as if conducting an invisible orchestra.

To Err is Human(-like): Effects of Robot Gesture on Perceived Anthropomorphism and Likability. Salem et al. (2013)

In an experiment with a humanoid robot, authors examined to what extent gesture affects anthropomorphic inferences made about the robot.

Specifically, they investigated the effects of the robot's hand and arm gestures on the perception of human likeness, likability of the robot, shared reality, and future contact intentions after interacting with the robot. For this purpose, the speech-accompanying nonverbal behaviors of the humanoid robot were manipulated in three experimental conditions: (1) no gesture, (2) congruent gesture, and (3) incongruent gesture. They hypothesized higher ratings on all dependent measures in the two multimodal (speech and gesture) conditions compared to the unimodal condition (speech only).

To Err is Human(-like): Effects of Robot Gesture on Perceived Anthropomorphism and Likability. Salem et al. (2013)

Findings:

The results confirm their predictions: when the robot used co-verbal gestures during interaction, it was anthropomorphized more, participants perceived it as more likable reported greater shared reality with it, and showed increased future contact intentions than when the robot gave instructions without gestures.

Surprisingly, this effect was particularly pronounced when the robot’s gestures were partly incongruent with speech, although this behavior negatively affected the participants’ task related performance. These findings show that communicative non-verbal behaviors in robotic systems affect anthropomorphic perceptions and the mental models humans form of a humanoid robot during interaction.

Mimicry and Imitation

Mimicry means the unconscious replication of the behavior of another person, whereas ***Imitation*** means the conscious replication of another's behavior (Genschow et al., 2017).

Social functions:

1. One is that it indirectly signals positive affect and liking for an interaction partner. If two people use the same gestures or adopt the same posture during a conversation, it is usually because they have established a positive relation in that interaction.
2. Similarly, when people's nonverbal cues are out of sync and not mirroring each other, you can sense that the communication is not running smoothly.
3. Mimicry as a subtle nonverbal cue can thus be a helpful signal to interpret, for instance, in the context of dating or job interviews.

Mimicry and Imitation

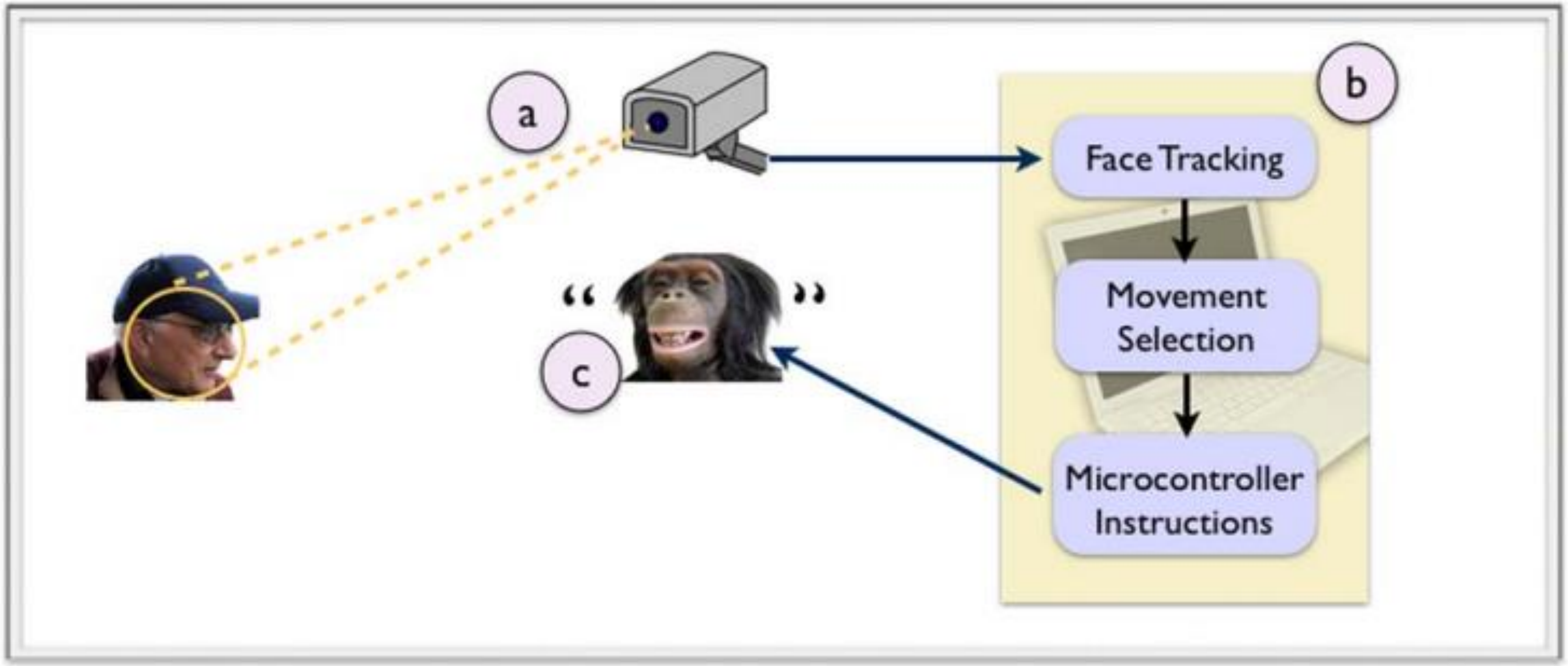
Development functions:

1. In early childhood development, mimicry and imitation provide a common way to learn new behaviors and culturally relevant social norms.
2. As adults, we can also use imitation to blend into our social and cultural surroundings, such as gesturing more emphatically when we are speaking Italian or visiting Italy.
3. Imitation and mimicry can be important ways of developing signs of in-group identity.

Mimicry and Imitation

Riek et al. (2009) had conducted an experiment to investigate how imitation by a robot affect people's perceptions of their conversation with it. The robot operated in one of three ways: *full head gesture mimicking*, *partial head gesture mimicking (nodding)*, and *non-mimicking (blinking)*.

The end-to-end system. (**a**) The video camera captures the user's head movements. (**b**) The OpenCV face tracker extracts the facial movements, the movement selector decides which way to move the robot, then microcontroller instructions are transmitted to the robot. (**c**) The robot moves



Mimicry and Imitation

They performed gesture analysis to see if any differences existed between groups and did find that men made significantly more gestures than women while interacting with the robot as they usually do in dyadic conversation.

Two participants, both in the partial mimicking (nodding) condition, co-nodded with the robot. In other words, the participant nodded, the robot nodded in response, and then the participant nodded to acknowledge the robot’s nod. This ability for the robot to influence the user’s behavior was something entirely unexpected.

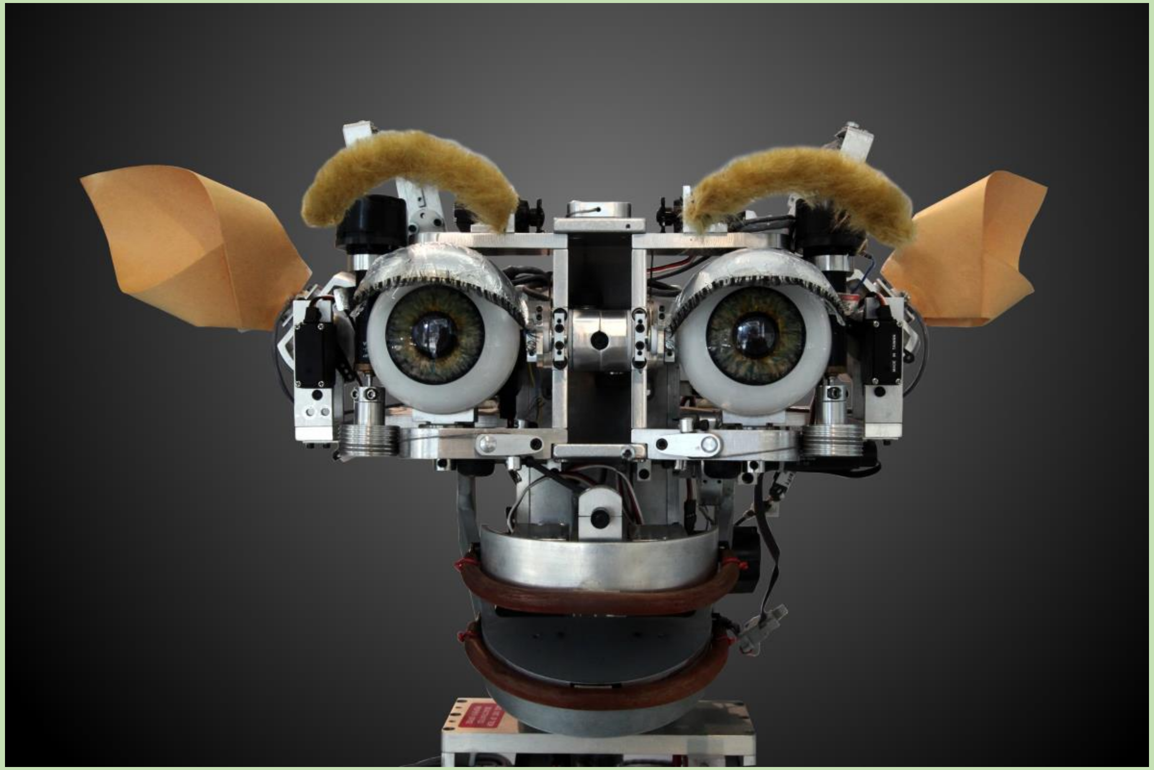
Mimicry and Imitation

Wills et al. (2016) examined the effect of socially contingent robot behaviors on a charity collection task. Manipulating only behavioral cues (maintaining the same verbal content), they showed that when the robot exhibits contingent behaviors consistent with those observable in humans, this resulted in a 32% increase in money collected over a non-reactive robot. These results suggest that apparent social agency on the part of the robot, even when subtle behavioral cues are used, can result in behavioral change on the part of the interacting human.

Robot Faces

Mechatronic Face

Kismet [5] is one of the earliest and most classic mechatronic expressive robot, with all features–such as eye lids, eyebrows, lips and ears–being physically implemented and controlled by electric motors.



Robot Face

Android faces have a larger number of mechatronic actuators controlling a flexible plastic skin. Examples are the Hanson robot faces, such as the Albert Huboor Joey Chaos robot heads [8] or Ishiguro’s androids [9]. Android heads, due to the number of actuators and the nonlinear interaction with the plastic skin, are typically more expressive than the abovementioned mechatronic heads. However, they are mechanically very complex and as such expensive to design, construct and maintain.

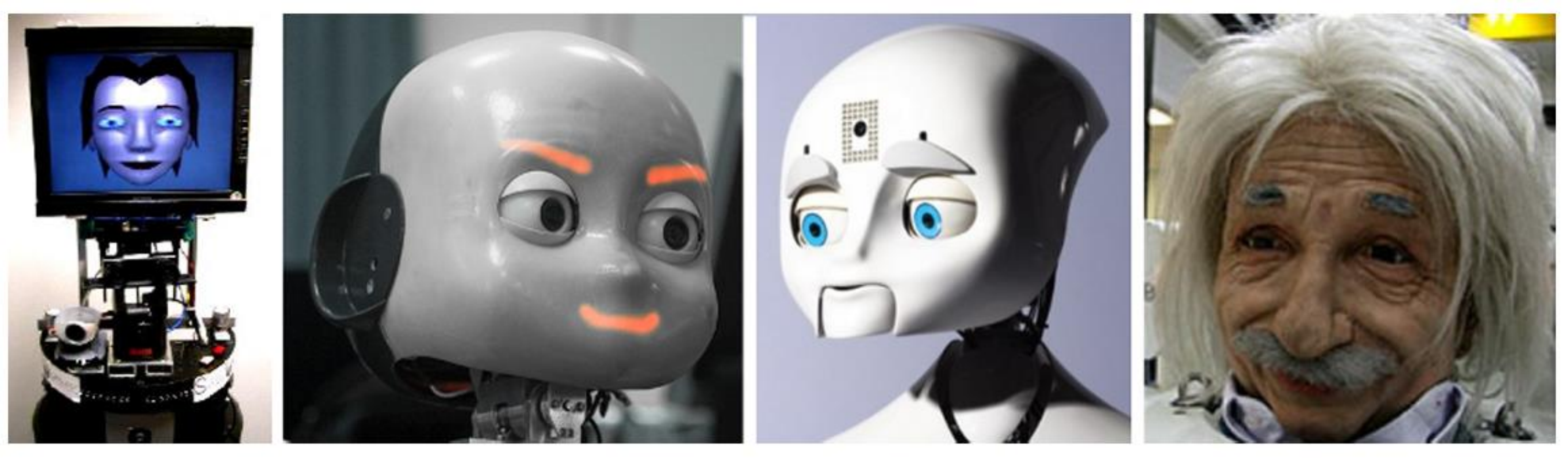
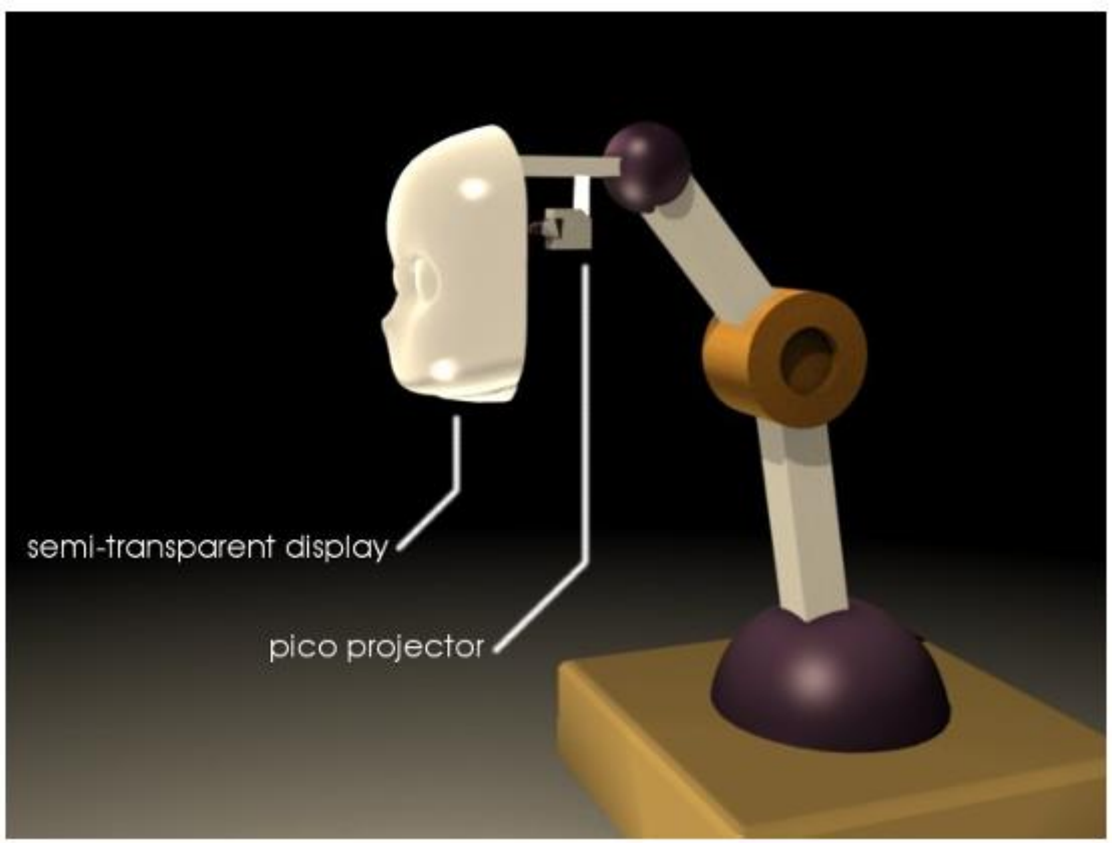


Figure: Some examples of contemporary robot heads: flat screen head GRACE, i-Cub mechatronic and LED head, Nexi MDS mechatronic robot head and Albert Hubo android robot head.

Robot Face

An alternative to mechatronic and android heads is to retro-project a robot face into a face-shaped translucent surface.

Delaunay et al. designed for an on-line animation retro projected on to a mask, which is not flat but has the shape of a human face. A complete HRI solution should support the head movements, so some mechanical parts were still required for proper interaction with the user.



Sketch of the robotic setup, the image is retro-projected into the opal face mask. The face is mounted onto an articulated neck. The current prototype described in this paper uses a ViewSonic XGA projector, however a Pico projector is to be used in a later version.

Touch

Touch is a nonverbal cue that is often involved in close interactions among people, such as those between friends or between caregivers and patients.

Touch can be deliberate (loved ones, between friends) or incidental (unintentional contact); functional (care giver and patient) or social (handshake, object handover and manipulation).

Functionalities of touch:

1. The affective touch hypothesis proposes that touch has an important effect on emotional well-being. Positive emotions evoked through gentle strokes (e.g., grooming) seem to mediate socializing behavior and appear to play a crucial role in autism. Furthermore, touch deprivation is linked to depression and violence. (Urakami et al. 2023)

Touch

3. On the other hand, cultural groups have different rules about who, how, and how often members can touch each other. Violations of these rules can have social and legal consequences, such as touch or body contact that is perceived as being inappropriate or unwanted.

4. Touch initiation is a dominant behavior that controls or directs the behavior of others. For instance, people perceived to be of higher status may initiate touch.

5. Touching also fulfills basic psychological needs. Hugging friends and family members or stroking the warm fur of animals creates positive feelings, is comforting, and reduces stress.

Touch

Yohanan and MacLean (2012) investigated the human's intent and expectations when displaying emotional touch. This touch was directed to a Haptic Creature robot, which was imagined by participants as their close pet.

Haptic Creature

Participant interacting with Haptic Creature during study

Findings:

1. They compiled a dictionary of plausible touch gestures for use in the study. They documented which gestures the human was more likely to use and for which specific emotions. *Given the stated relationship with the Haptic Creature, participants gravitated to less aggressive, more affectionate gestures.*

Touch

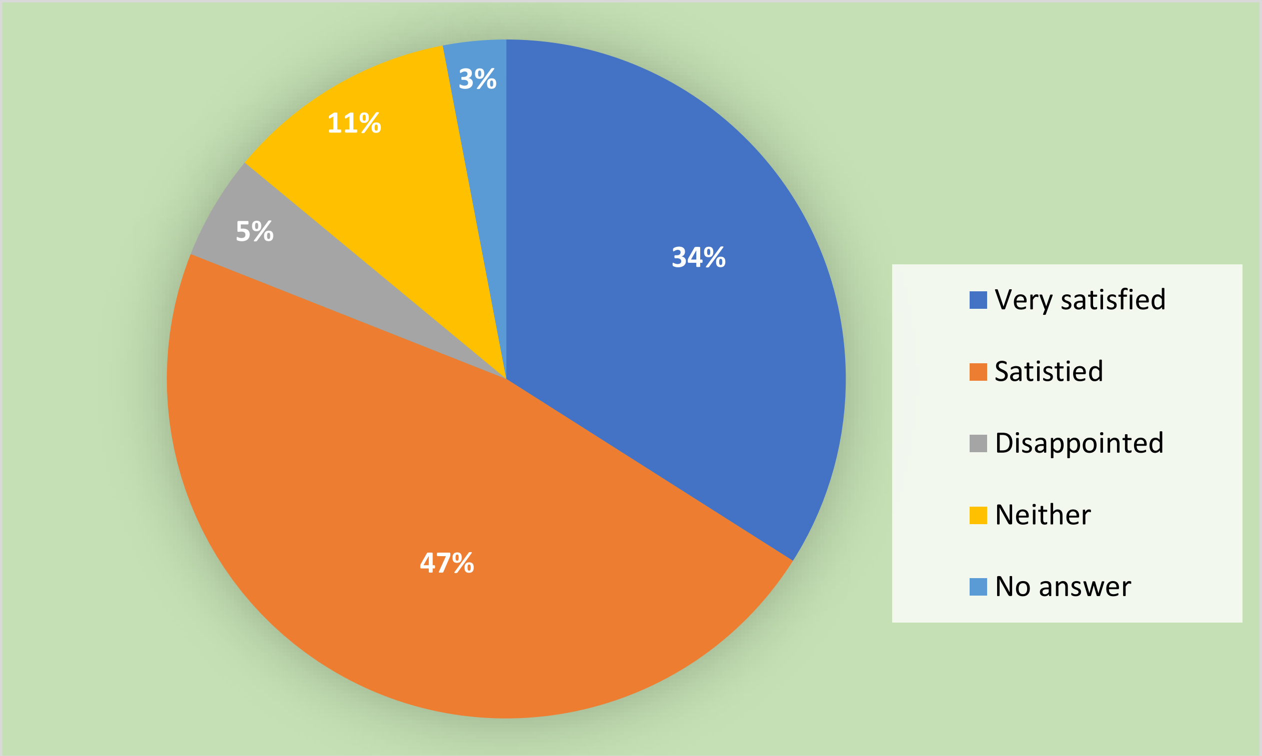
2. They presented the physical properties of the touch interaction, which included common points of contact as well as duration and intensity of gestures. Results showed the *human's fingers and hands to dominate repetitious gestures*, while *forearms and chest were included in sustained gestures*. For the Haptic Creature, *the back was the major point of contact from the human*, while its aft sides were included in many sustained gestures.

3. They categorized human's intent when communicating emotion through touch to the robot. The categories are protective, comforting, restful, affectionate, and playful. *Participants demonstrated an expectation that the Haptic Creature provide an emotional response like the ones they were communicating*.

Touch

Shibata et al. (2011) examined people living with Paro, a seal robot, and how they interact with it. Questionnaires were sent along with Paro to the owners who could voluntarily choose to respond. Eighty-five responses were obtained and analyzed in terms of *gender*, *preference for animals*, and *pet ownership experience*.


Paro, the seal robot



Touch

Findings:

1. The results showed that for females, besides "can touch and hug," "tactile texture" is also an important reason for keeping Paro. This finding suggests that pet type robots designed to interact with people need to have good tactile texture to be accepted by females.

2. In addition, owners' preference for animals seems to influence how they pet Paro. Owners who have a dislike or no preference for animals tend to touch Paro differently than do those who like animals, seemingly because of their different playing styles. Thus, the difference may be due to different playing styles based on owners' previous experience in interacting with animals.

Touch

Findings:

3. Moreover, owners who have pet ownership experience tend to talk to Paro more when they wake up and go to sleep compared to owners who have never kept pets. This suggests that word recognition functions such as the ability to recognize “Praise words” and “Greetings” are demanded for pet type robots, especially by people with pet ownership experience, and may be important to be accepted for pet type robots, which interact with people.

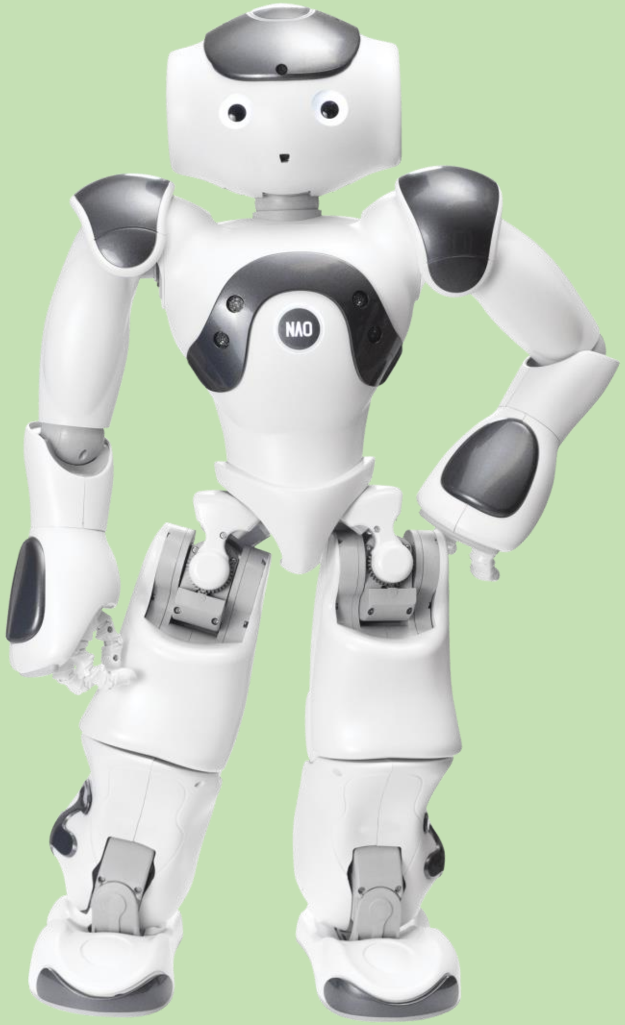
4. Furthermore, owners who have pet ownership experience demand a “collar” as an accessory for Paro more than do those who have no previous experience of keeping pets.

Touch

Wullenkord et al. (2016) examined how Contact Type (Actual vs. Imagined) and Contact Modality (Look vs. Touch) with a NAO robot would impact attitudes toward NAO compared to a no-contact control condition.

Findings:

1. Results showed that nearly any type of contact effectively reduced negative emotions compared to the control condition.
2. However, for participants with preexisting negative emotions toward robots, contact sometimes produced more negative attitudes.



Posture and Movement

Posture refers to a special position of the body or the way that a person holds their body. The four basic body postures are *standing*, *sitting*, *squatting*, *and lying down*.

There are many variations in postures, but a basic distinction can be made between *dynamic postures* (when the body is in motion, e.g., walking) and *static postures* (when the body is not moving, e.g., sleeping or standing).

Functions of Posture:

1. Along with facial expression, postures can be used to interpret a person’s emotional state. Slow movements, drooping shoulders, and lethargic gestures all suggest a downcast state of mind, whereas fast movements and an upright bearing are signs of a positive attitude.

Posture and Movement

2. Posture can signal attention, engagement, and attraction in an interaction between humans. If two people are sitting with their knees toward each other, it shows willing engagement, whereas if one person is turned partly away from the other, it can show a desire to discontinue the interaction.

3. Position and posture can communicate the attitude and relationship status of a sender. For example, Mehrabian et al. (1969) found that when communicators who were women adopted an open posture, they were interpreted as having a more positive attitude than when they adopted a closed posture.

These types of postural cues are particularly important when a person's face is not visible, but they can also provide additional cues to a person's state of mind even when we can see the person's facial expression.

Researchers have found that people can interpret these types of nonverbal cues not only when they see the whole body of the person but also in minimalist light dot displays that depict a person's movements (Alaerts et al., 2011).

Posture and Movement

Is posture can be identified by human?

Beck et al. (2012) confirmed that that it is possible to accurately identify emotions displayed by body language alone and found that the physical appearance of the body does not affect the identification of an emotion.

Expressive key poses can be accurately interpreted and that, as for humans, head position is highly expressive. It was found that moving the head down leads to decreased arousal (defines the level of energy), valence (defines whether a stimulus is positive or negative) and stance (defines whether a stimulus is approachable) whereas moving the head up increases these three dimensions.

Posture and Movement

Emotional body language is an appropriate medium for expressing emotions even with simplified bodies such as the one of a humanoid robot. Moreover, the study of emotional body language can benefit from the use of humanoid robots such as Nao as they can display rich emotional body language, similar to the one displayed by humans, without being confounded by other modalities such as facial expressions.

Posture and Movement

Xu et al. (2014) tried to find whether

(i) Can participants differentiate between positive and negative robot mood expressed in gestures during an interaction scenario, rather than in a pure recognition task?

(ii) Can mood expressed by a robot induce mood contagion effects in human observers?

(iii) Can the mood expression of a robot influence the performance of a human in an interaction task?

For example, it is well known that mood can transfer between persons and has specific effects on behavior (Neumann et al. 2000) and it is useful to gain insights into the effects and possible transfer of mood from a robot to an individual.

Findings: They showed that people were not only able to interpret the affective body postures of robots, but also that *they adopted the emotions they thought the robots were showing*.

Interaction rhythm and timing

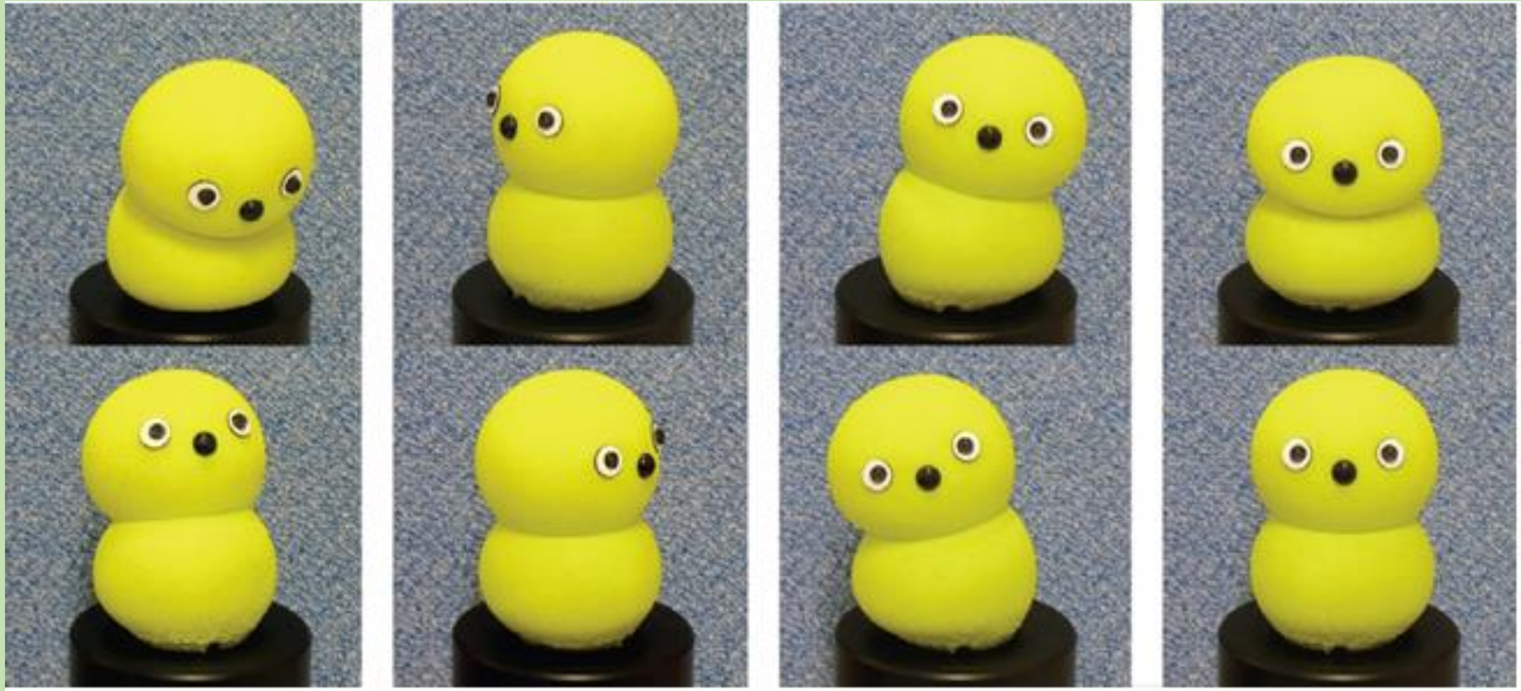
The “rhythmicity” and “synchrony” of an interaction provide a largely unconscious but crucial component of human communication.

In verbal communication, we refer to this as turn-taking among interaction partners. Nonverbal cues (e.g., gaze, gesture) can support this turn-taking by guiding attention to the appropriate interaction partner or signifying the end of a turn. Establishing synchronized temporal patterns of interaction can further scaffold the communicative and collaborative success of an interaction.

Interaction rhythm and timing

Michalowski et al. (2007) showed that a robot that is rhythmically entrained with a human interaction partner is considered more lifelike than a robot that is behaving rhythmically but is not synched with the human.

They also showed that people are more likely to interact for a longer time with a rhythmically entrained dancing robot.



Keepon’s four degrees of freedom (nodding, panning, rocking, and bobbing).

https://www.youtube.com/watch?v=nPdP1jBfxzo

Robot perception of nonverbal cues

How can robot understand the nonverbal cues from human?

Is the human really engaged with the robot through nonverbal interaction?

Robot perception of nonverbal cues

Standard pattern-recognition techniques are used to allow robots to perceive and identify human nonverbal cues.

Typical systems use cameras, depth cameras, or sensors carried by the user to record a time series of data.

Software could be written to recognize a limited number of gestures; it is instead typical to use *machine learning* as the system to be trained to recognize gestures and other nonverbal cues.

Robot perception of nonverbal cues

Engagement is“ the process by which two (or more)participants establish, maintain and end their perceived connection during interactions they jointly undertake”.

Based on a study of the engagement process between humans, Rich et al. (2010) had developed and implemented an initial computational model for recognizing engagement between a human and a humanoid robot. Their model contained recognizers for four types of connection events involving gesture and speech: directed gaze, mutual facial gaze, conversational adjacency pairs and backchannels. They used Robot Operating System (ROS) framework.

Generating nonverbal cues in robots

HRI poses challenges for perception and generation of nonverbal cues because all this must be done in real-time.

Challenges related to generation of nonverbal cues in robot:

1. The nonverbal cues need to be contingent on the interaction (cause and effect should be similar); if the user snaps her fingers, the robot needs to blink immediately.

2. Nonverbal cues also need to be coordinated with each other and with other cues, including verbal interaction, both in terms of the semantic meaning and the timing of execution.

Generating nonverbal cues in robots

- Animation framework: The simplest and most frequently used approach is to generate motions with an animation framework. That is, a robot designer will typically control each of the joint angles of a robot to set a posture for it; this is called a “key frame”. After the designer prepares multiple key frames, the system interpolates the postures between them and generates smooth motions for the robot.
- Motion-capture systems can be used to record a timed series of precise human motions, which can then be replicated in robots.

Generating nonverbal cues in robots

- Cognitive mechanisms for robots: Another approach to achieving natural behavior in robots is to endow the robot with artificial cognition, which is an artificial equivalent of natural cognition. The expectation is that natural interaction behavior will emerge in the robot when it is controlled by artificial cognitive mechanisms. So instead of hacking the robot's nonverbal behavior, a constructivist approach is used.

Liu et al. (2013) developed a computational model for a robot that balances two factors, *understandability* and *social appropriateness*. It enables a robot to refrain from exhibiting impolite pointing gestures while still keeping its deictic interaction understandable.

Reading Material

- Jacqueline Urakami and Katie Seaborn. 2023. Nonverbal Cues in Human–Robot Interaction: A Communication Studies Perspective. J. Hum.-Robot Interact. 12, 2, Article 22 (June 2023), 21 pages. https://doi.org/10.1145/3570169