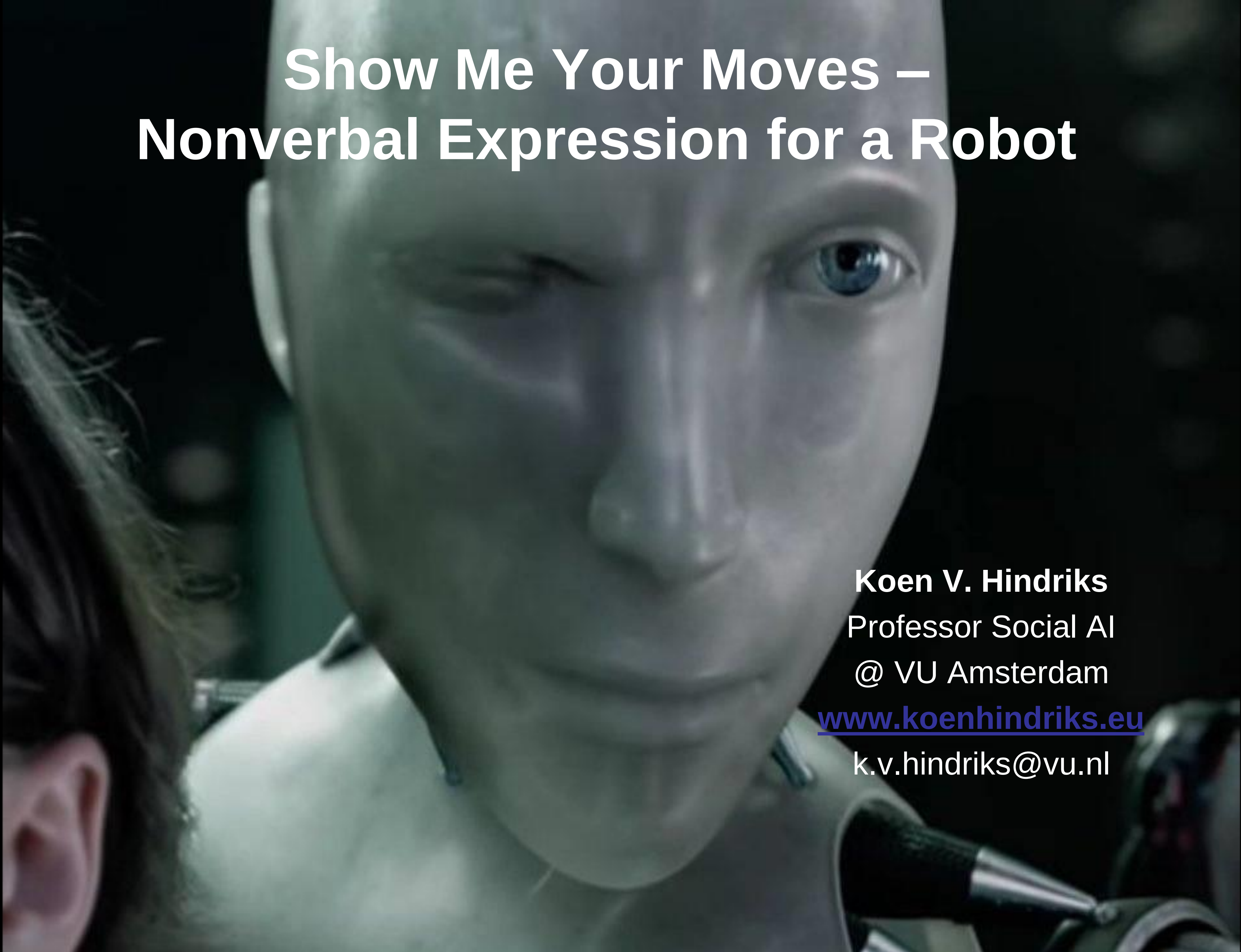
Show Me Your Moves –
Nonverbal Expression for a Robot



Koen V. Hindriks
Professor Social AI
@ VU Amsterdam
www.koenhindriks.eu
k.v.hindriks@vu.nl

Goal & Learning Objectives

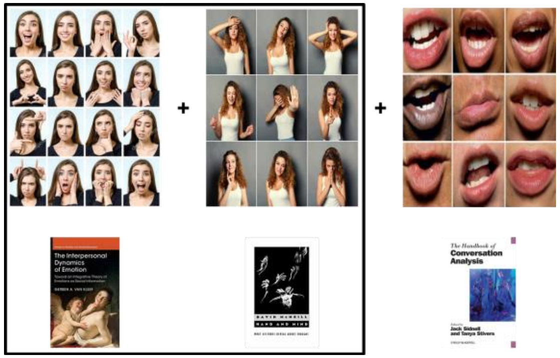
Goal:

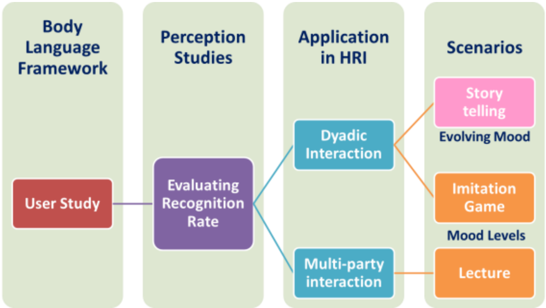
- The goal of this lecture is to introduce you to nonverbal robot communication, in particular to the expression of affect.

Learning Objectives:

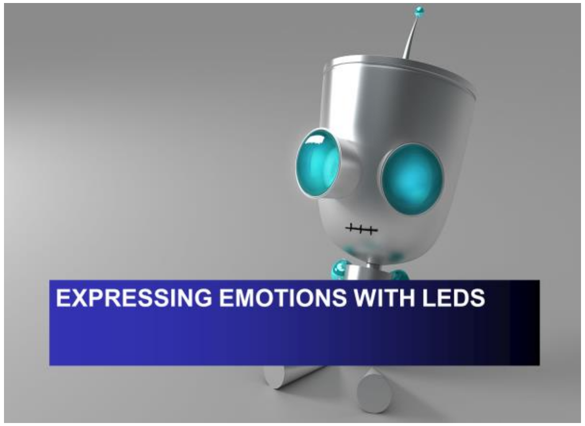
- Able to explain what nonverbal communication is
- Understand some modalities of HH non-verbal communication
- Able to explain what affect is
- Able to explain which parameters influence affect expression

Overview





Body Language Framework
Perception Studies
Application in HRI
Scenarios
Story telling
Evolving Mood
Dyadic Interaction
User Study
Evaluating Recognition Rate
Imitation Game
Mood Levels
Multi-party interaction
Lecture

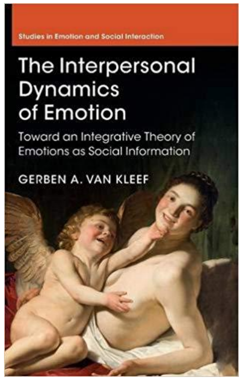


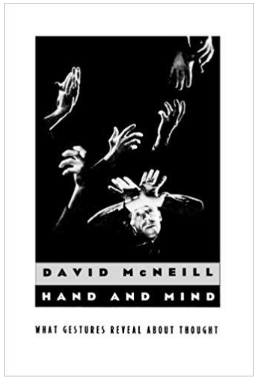
EXPRESSING EMOTIONS WITH LEDS

Language of Social Interaction

+

+



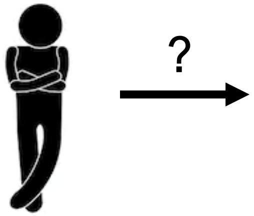


Non-verbal communication

- Nonverbal *modalities*: different ways in which a social agent communicates information without words → social cues

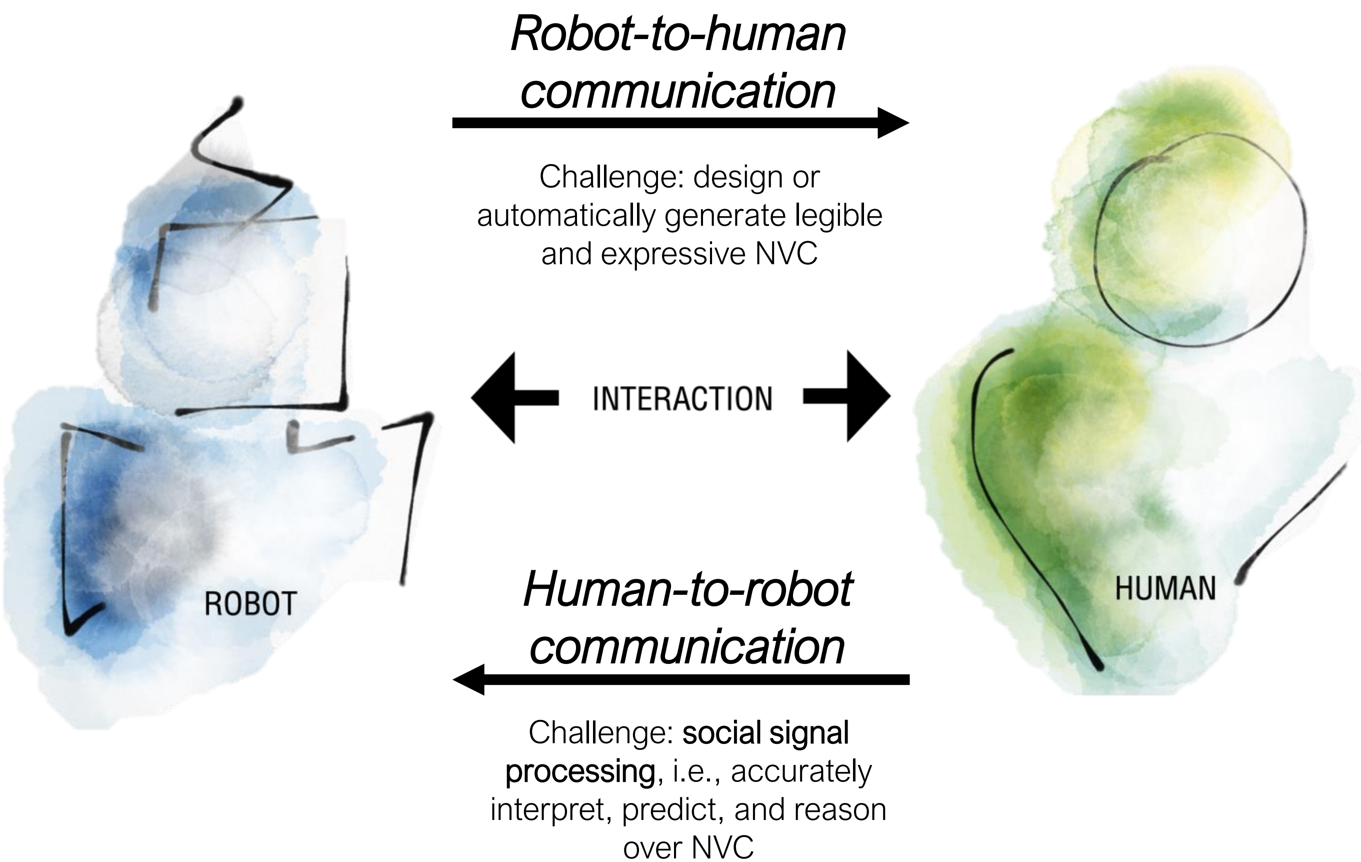
- Embodied interaction: physically present, but (for humans) can also be mediated, only social presence

- Challenges for robot NVC
  - Translate principles of human NVC to robots
  - Different embodiment than humans
  - Unique modalities (e.g., lights, sound)



Sphero robot

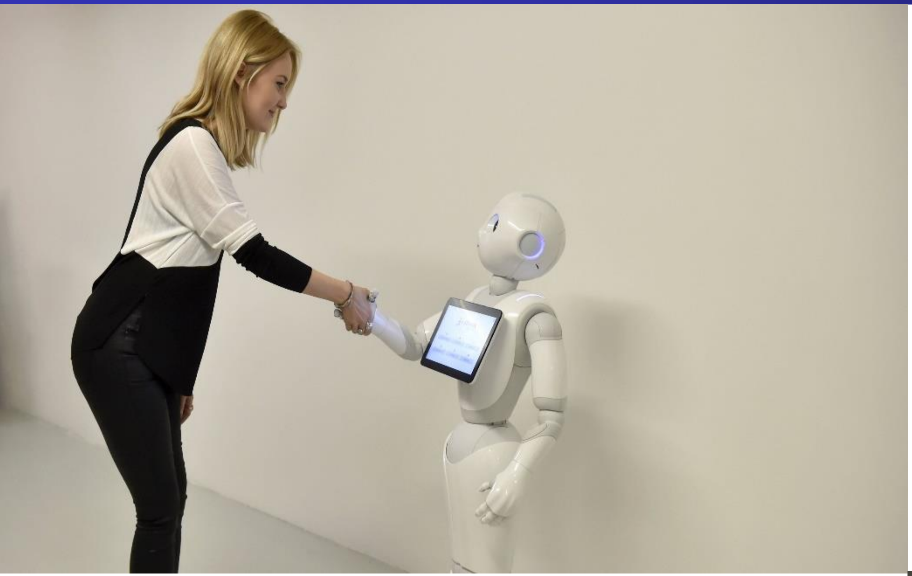
Communication is always a two-way street



Kim Baraka - SIR '21 - Nonverbal communication

MOTIVATION

A New Interface: Emotion



Patient Acceptance

Variable	Score (0-10)
Attitude towards technology	7,2
Facilitating conditions	8,2
Anxiety	1,1
Perceived sociability	6,1
Social influence	5,4
Perceived ease of use	7,6
Social presence	5,2
Perceived enjoyment	6,2
Trust	6,4

- Lack of responsiveness to additional explanations of patients.
- **Emotional responses** of patients on being questioned about quality of life.
- Robot does not recognize patient in repeated encounters.
- Loss of attention, due to 'confusion' about question

Key Message

nonverbal communication plays a dominant role in social interactions

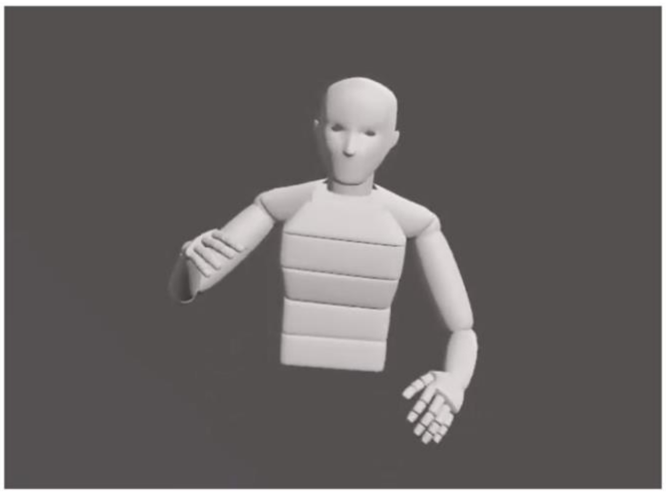
important to explore how nonverbal robot behavior influences humans

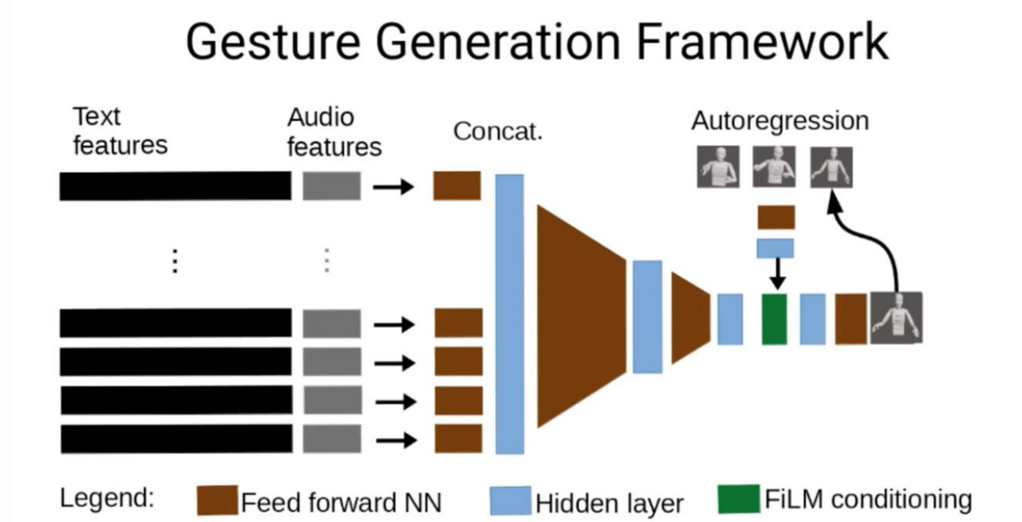
Nonverbal Communication

- **Vocalics:** pitch, volume, stress, backchanneling, word like sounds, etc.
- **Kinesics**: e.g. body movements, gestures, postures, facial expressions, eye gaze
- **Proxemics**: the conscious or unconscious setting of distances between various objects, agents, and oneself
- **Haptics**: robot touch, e.g. ‘pushing buttons’, hand shaking
- **Chronemics**: tempo of human interaction and the pace at which we expect communication to occur

Burgoon et al. 2022, Nonverbal Communication, Routledge

Automatic generation of speech-accompanying gestures





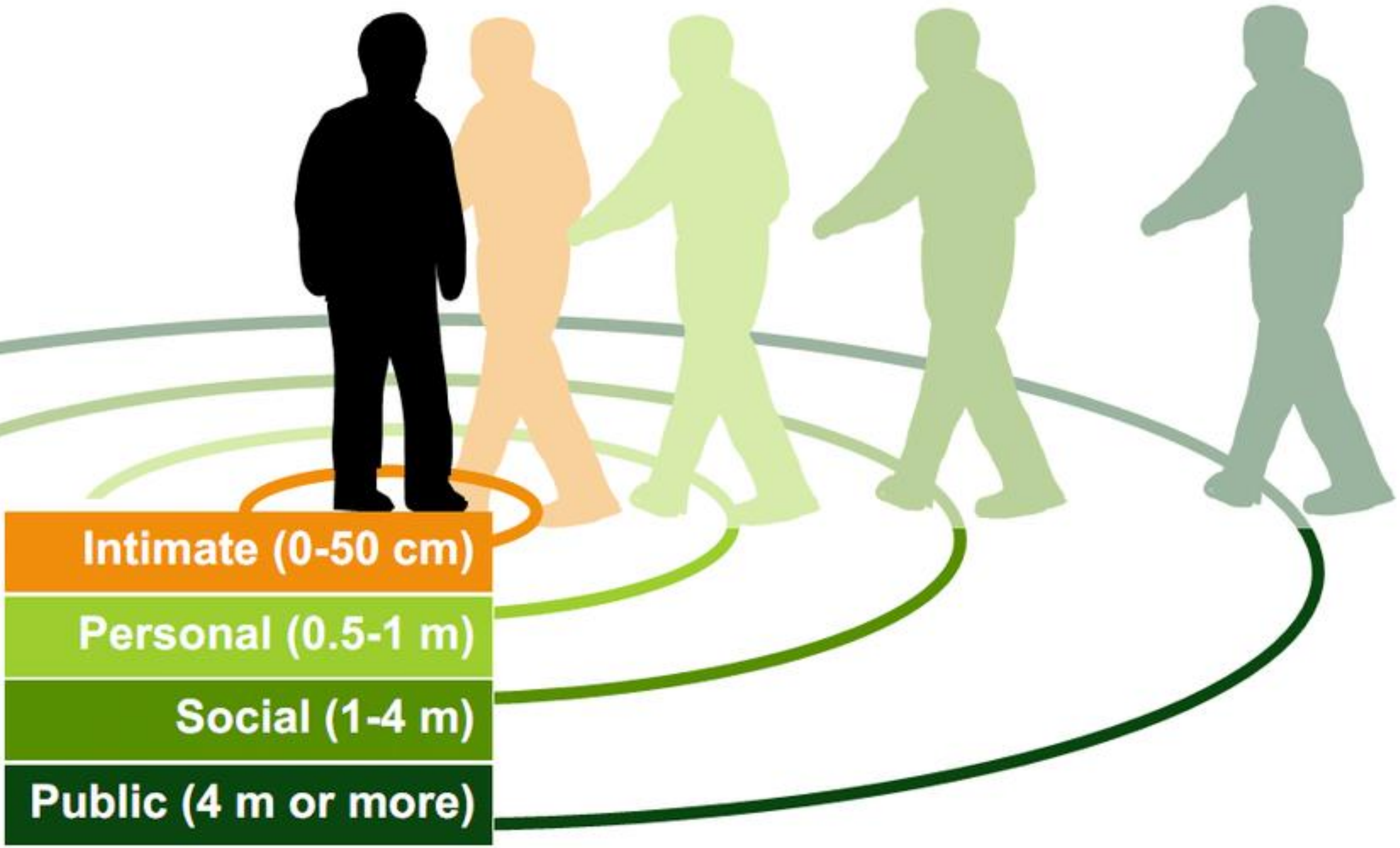
Gesticulator: A framework for semantically-aware speech-driven gesture generation (Kucherenko et al. 2020)

Gaze

- Gaze = where and how one looks
- People are uniquely sensitive to gaze – important to get it right on robots
- For non-anthropomorphic robots, gaze doesn’t necessarily have to involve realistic-looking eyes
- Three types of gaze
  - Mutual gaze (eye contact)
  - Deictic (“pointing” with your eyes)
  - Joint attention
- Types of eye movement
  - Fixation
  - Saccades
  - Smooth pursuits

JIBO “looking” at a cup

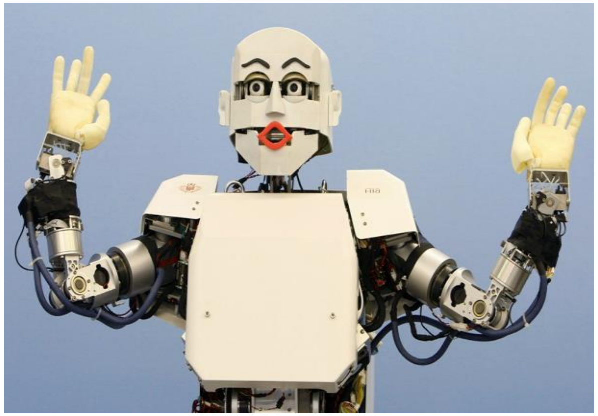
Hall’s Proxemic Zones

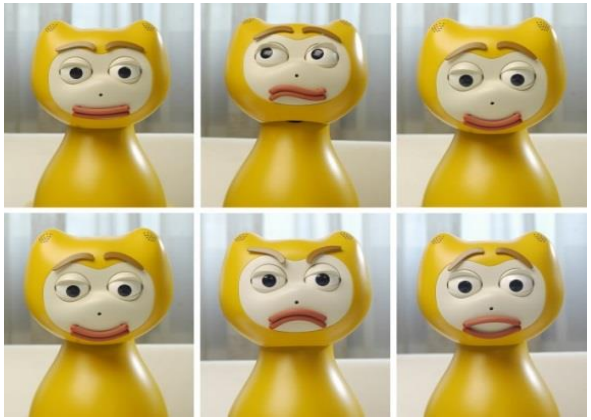


Functions of NVC

- Impression management
- Expressing emotions
- Expressing closeness
- Status
- Coordinating interaction
- Influencing others

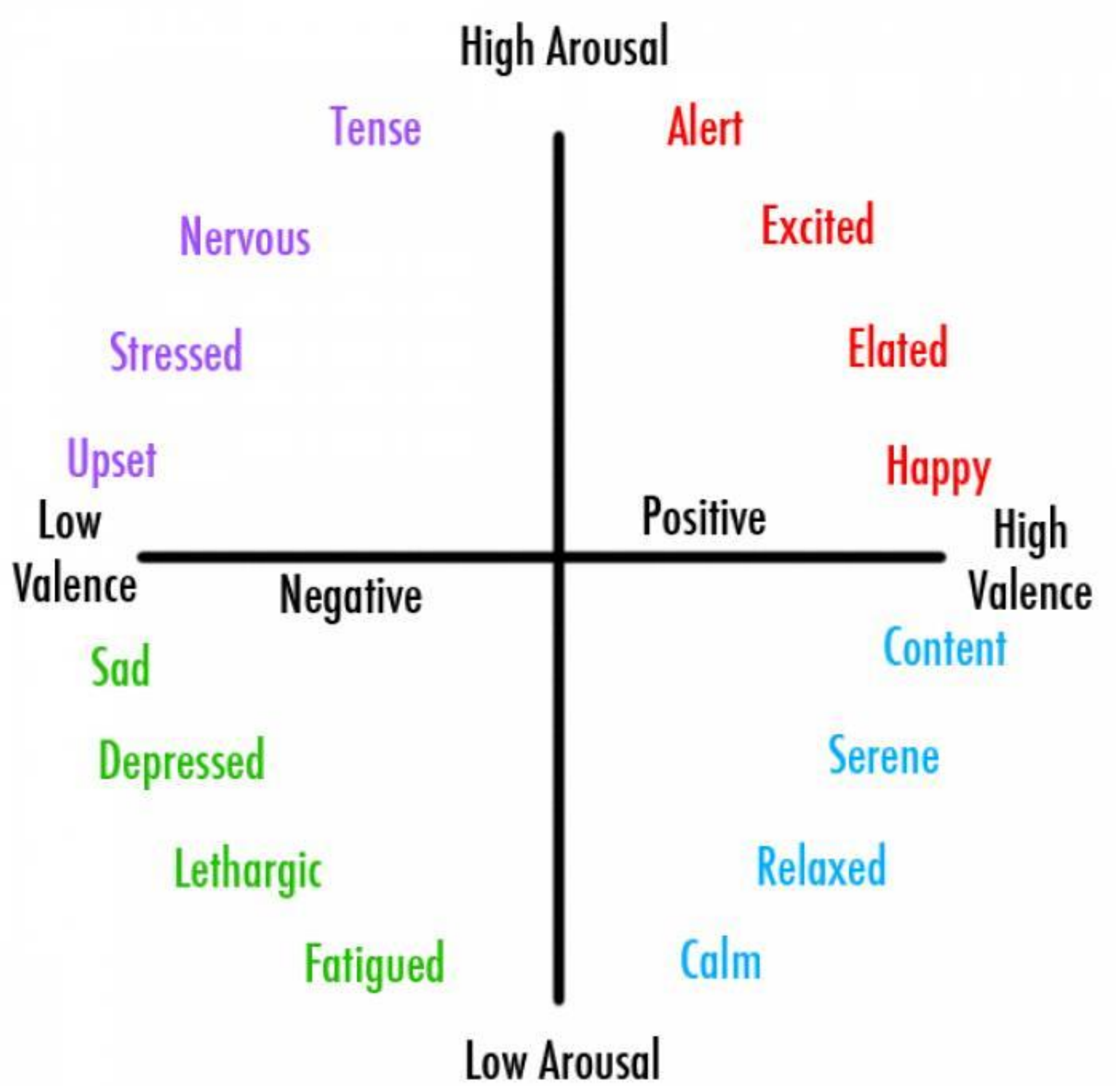
Nonverbal Affective Expression

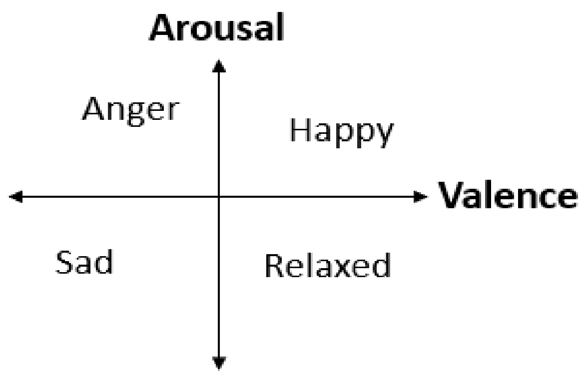




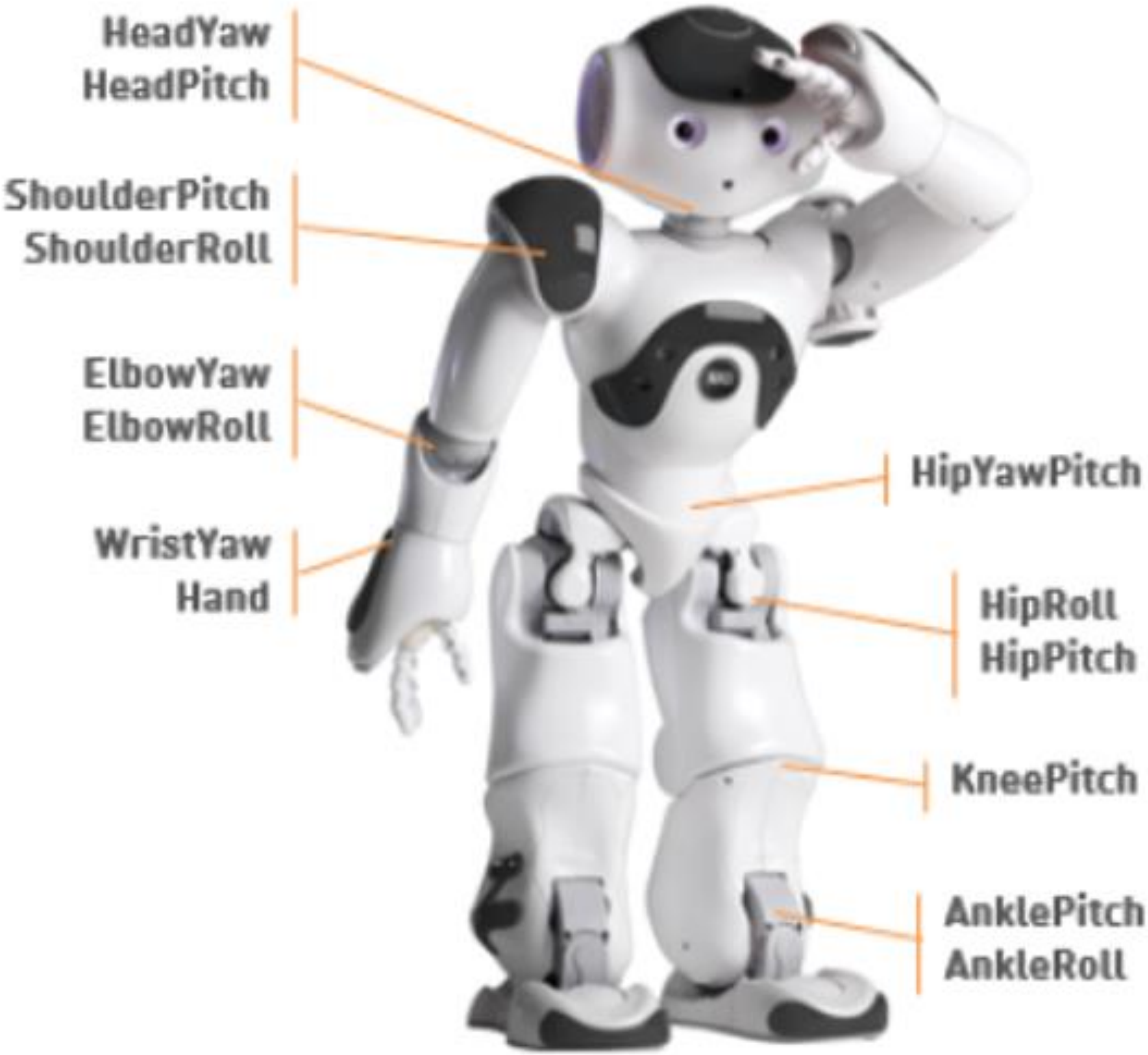
- Facial expression
- Nonverbal utterance
  - Voice/paralanguage;
  - Sound; Music
- Static posture/gesture; body movements
- Lights; LEDs; Color

What is an Emotion?

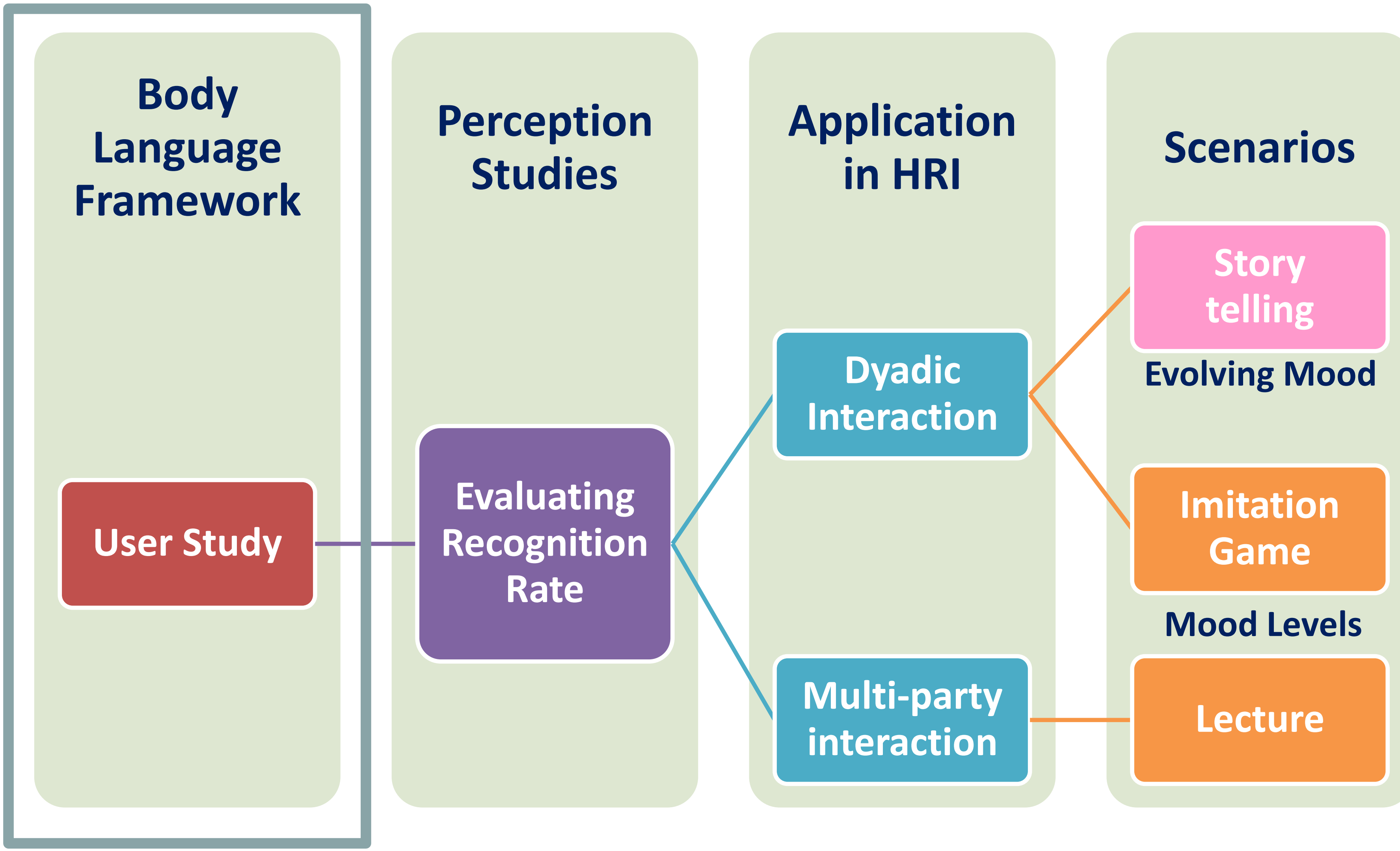




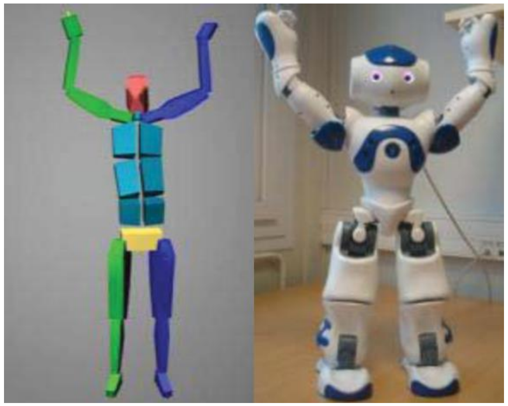
James A Russell. *A circumplex model of affect*, 1980.


HeadYaw
HeadPitch
ShoulderPitch
ShoulderRoll
ElbowYaw
ElbowRoll
WristYaw
Hand
HipYawPitch
HipRoll
HipPitch
KneePitch
AnklePitch
AnkleRoll

Overview


Body Language Framework
Perception Studies
Application in HRI
Scenarios
User Study
Evaluating Recognition Rate
Dyadic Interaction
Multi-party interaction
Story telling
Evolving Mood
Imitation Game
Mood Levels
Lecture

Emotional Body Language – Challenges



happy

surprise

angry

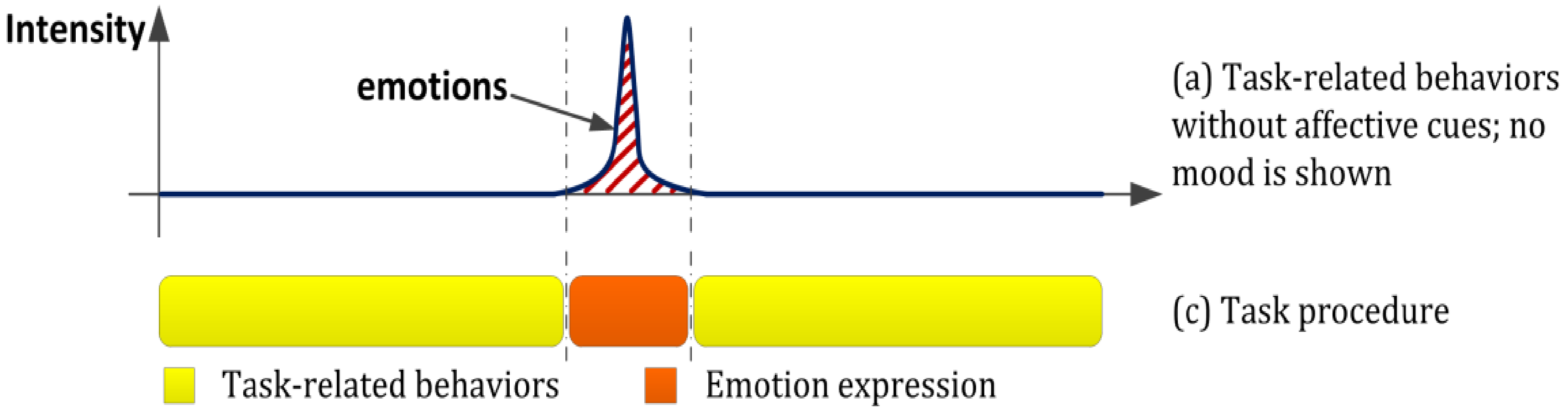
fear

sad

Expression through body movement:

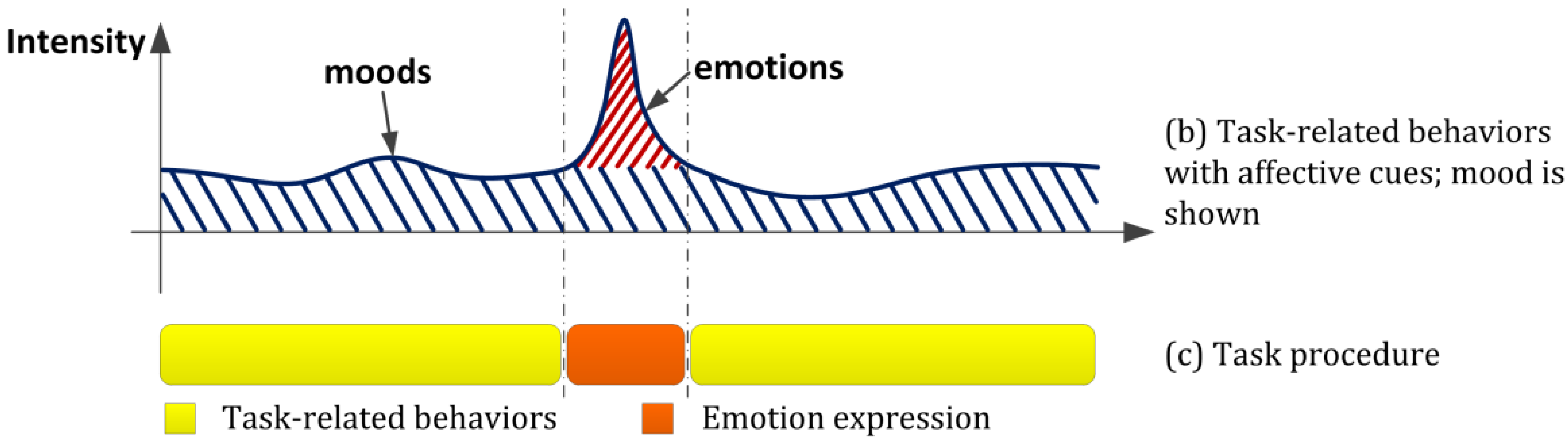
- Explicit body actions dedicated to communicating emotions; no other functions
- Interfere with task-related or functional behaviors
- Difficult to use during task execution

Emotional Body Language – Challenges



- Emotion expressions usually rise and decay quickly; they do not extend over time
- Affective state is not visible in between emotion expressions

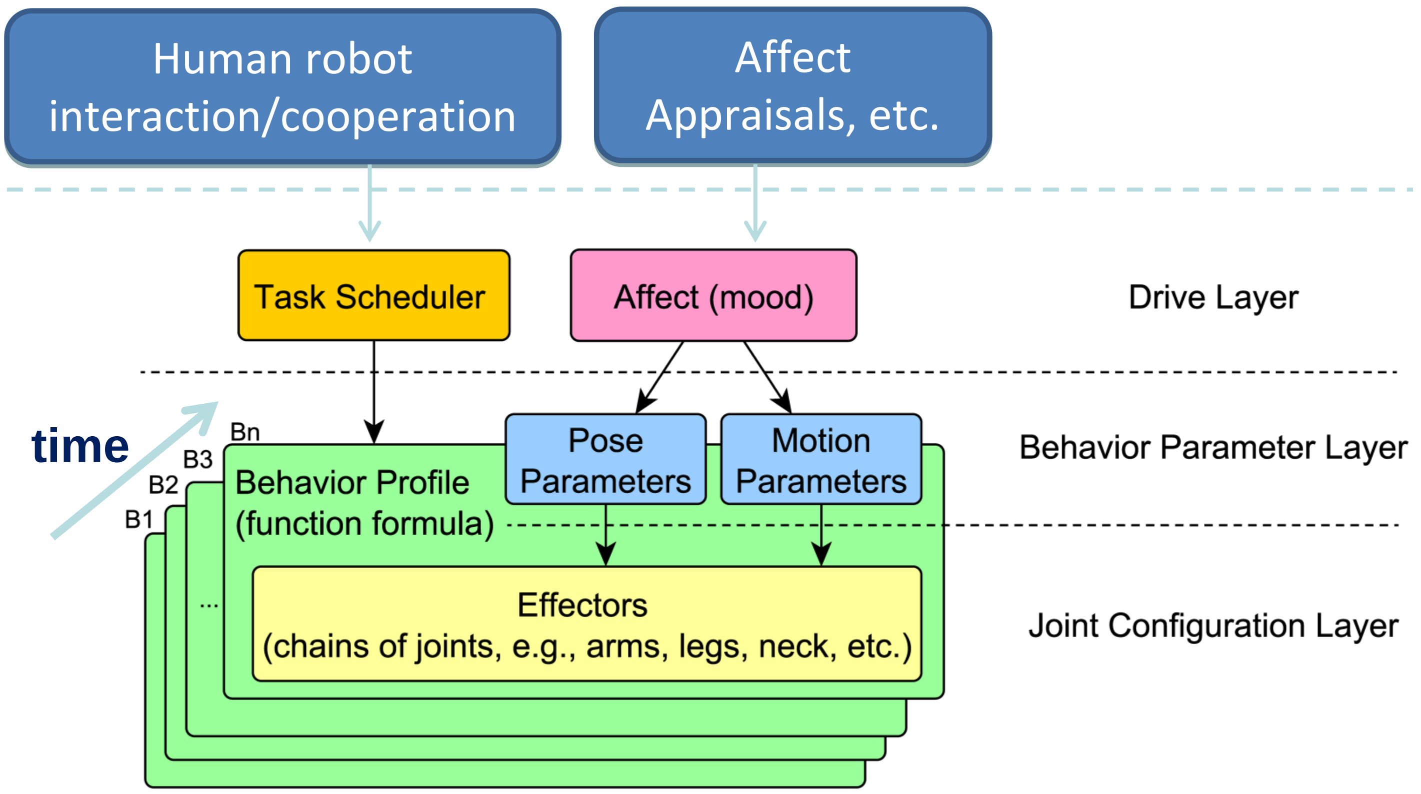
Emotional Body Language – Challenges



- Affect (mood) can be expressed simultaneously with task behaviors
- Mood expression indicates robots’ affect continuously over time

Parameterized Behavior Model

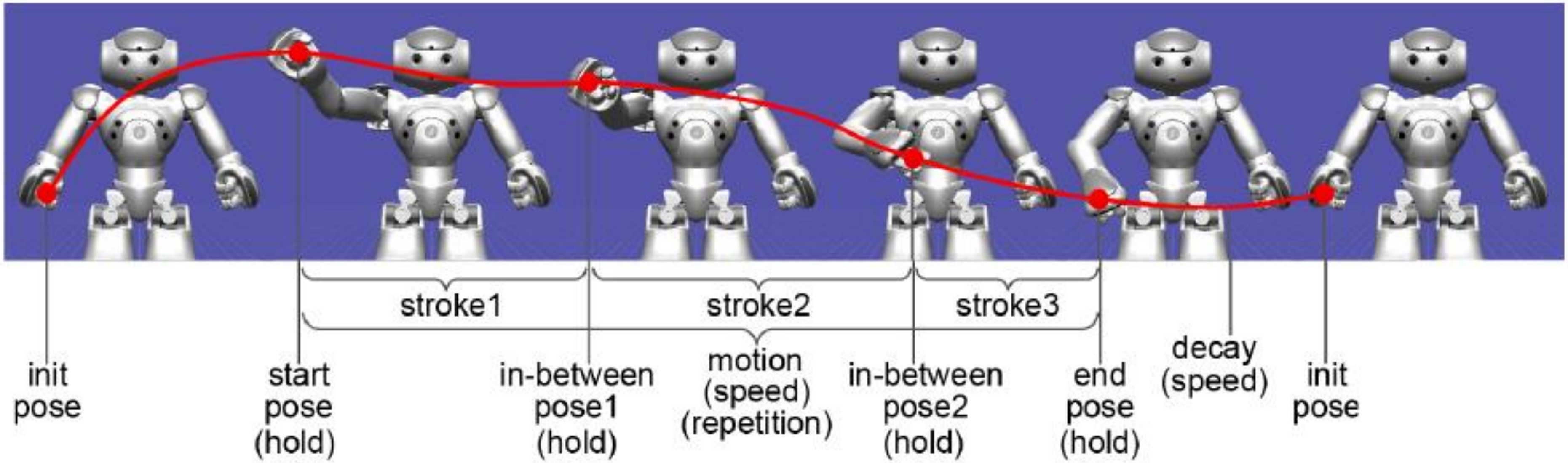
Behavior appearance varied through parameter modulation, to not interfere with functions



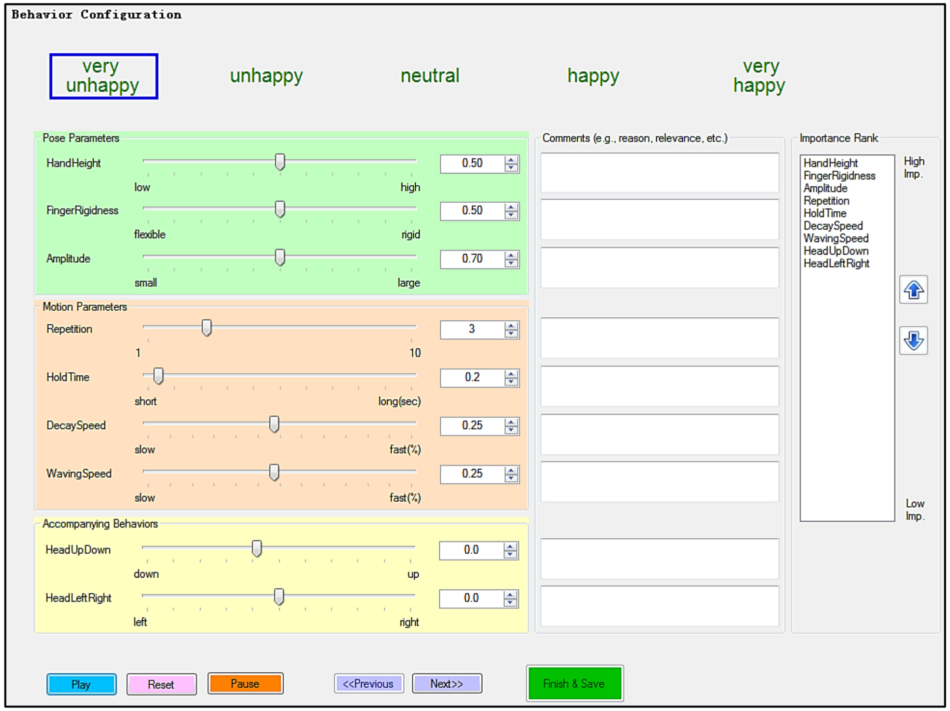
Parameter Space

Parameter	Negative	Positive
Hand-Height	low	high
Finger-Rigidness	bent	straight
Amplitude	small	large
Repetition	low	high
Hold-Time	long	short
Decay-Speed	slow	fast
Waving-Speed	slow	fast
Head-Up-Down	lower	raise
Head-Left-Right	look away	look at you

Dynamics: Example



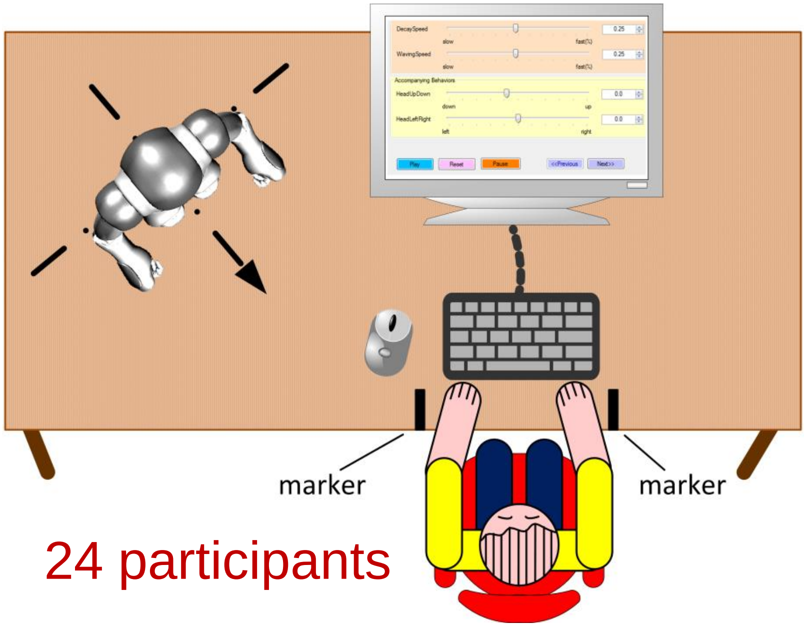
Fixing the Parameters – User Study



RQ1. Which parameters correlate (more) with mood (valence)?

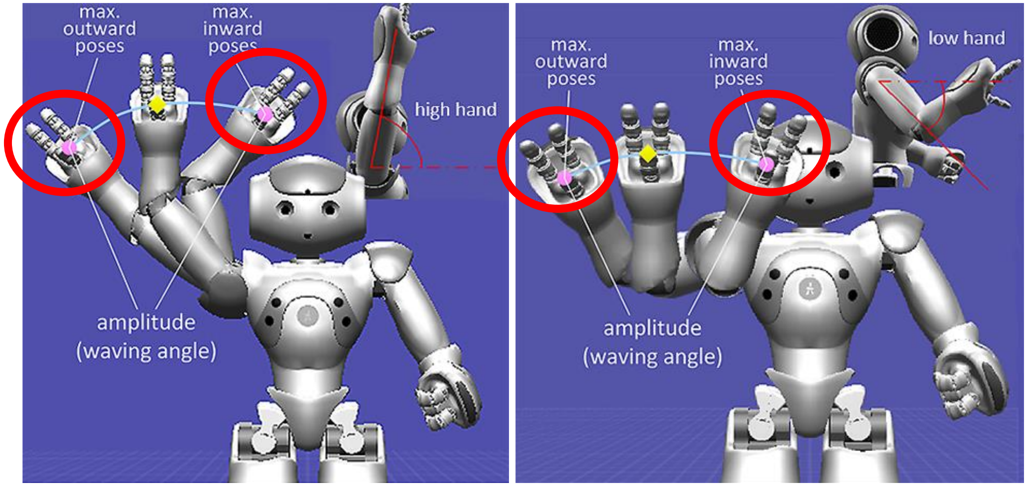
RQ2. What values should those parameters be for a certain mood level (five levels)?

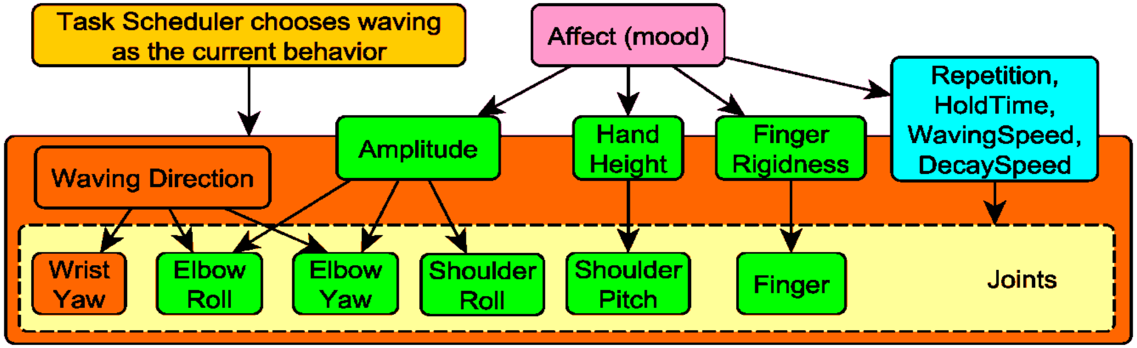
Combo-Orders	Behavior Orders	Mood Orders
1	Pointing → Waving	Negative → Positive
2	Pointing → Waving	Positive → Negative
3	Waving → Pointing	Negative → Positive
4	Waving → Pointing	Positive → Negative



24 participants

Pose Parameters - Waving

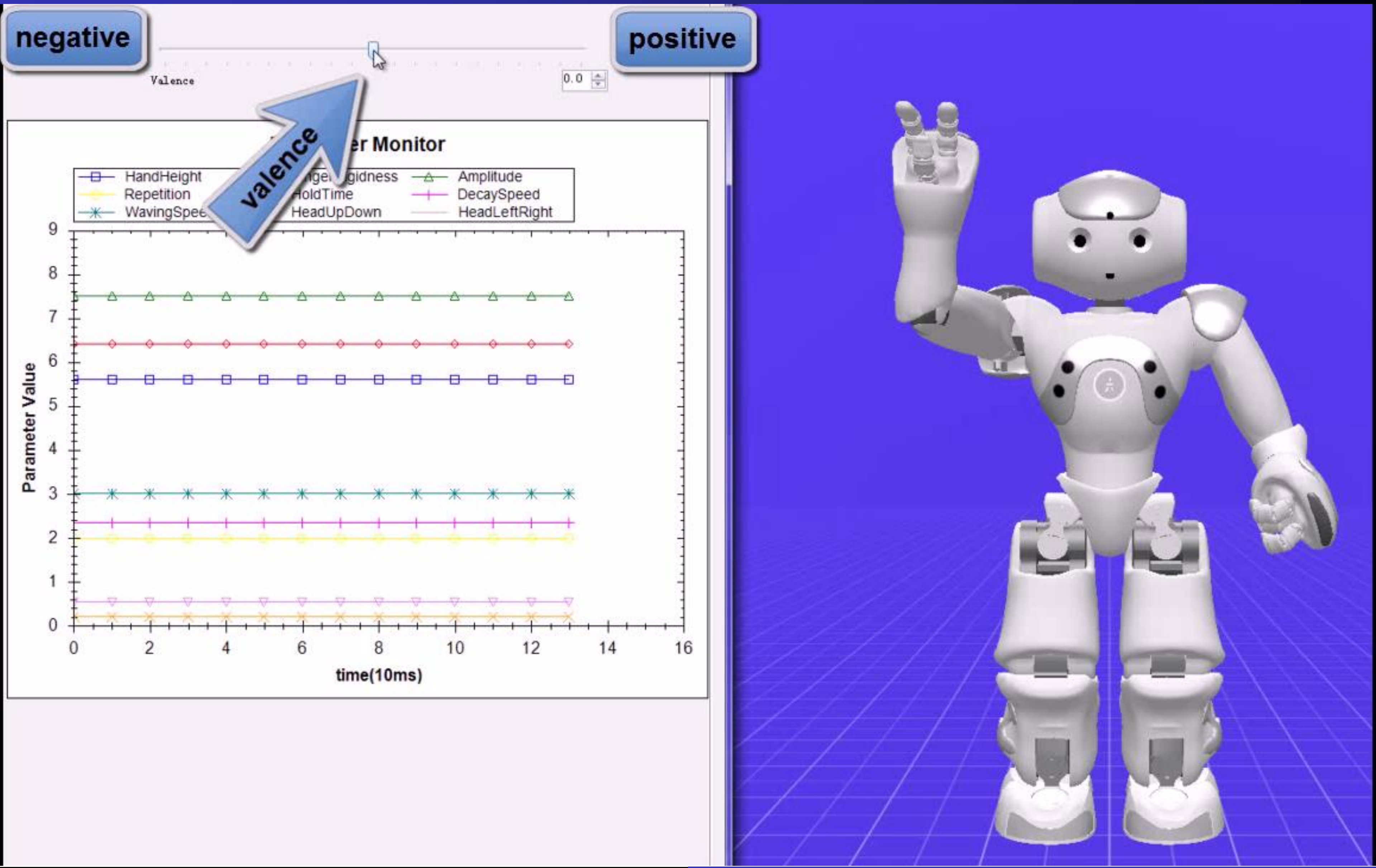




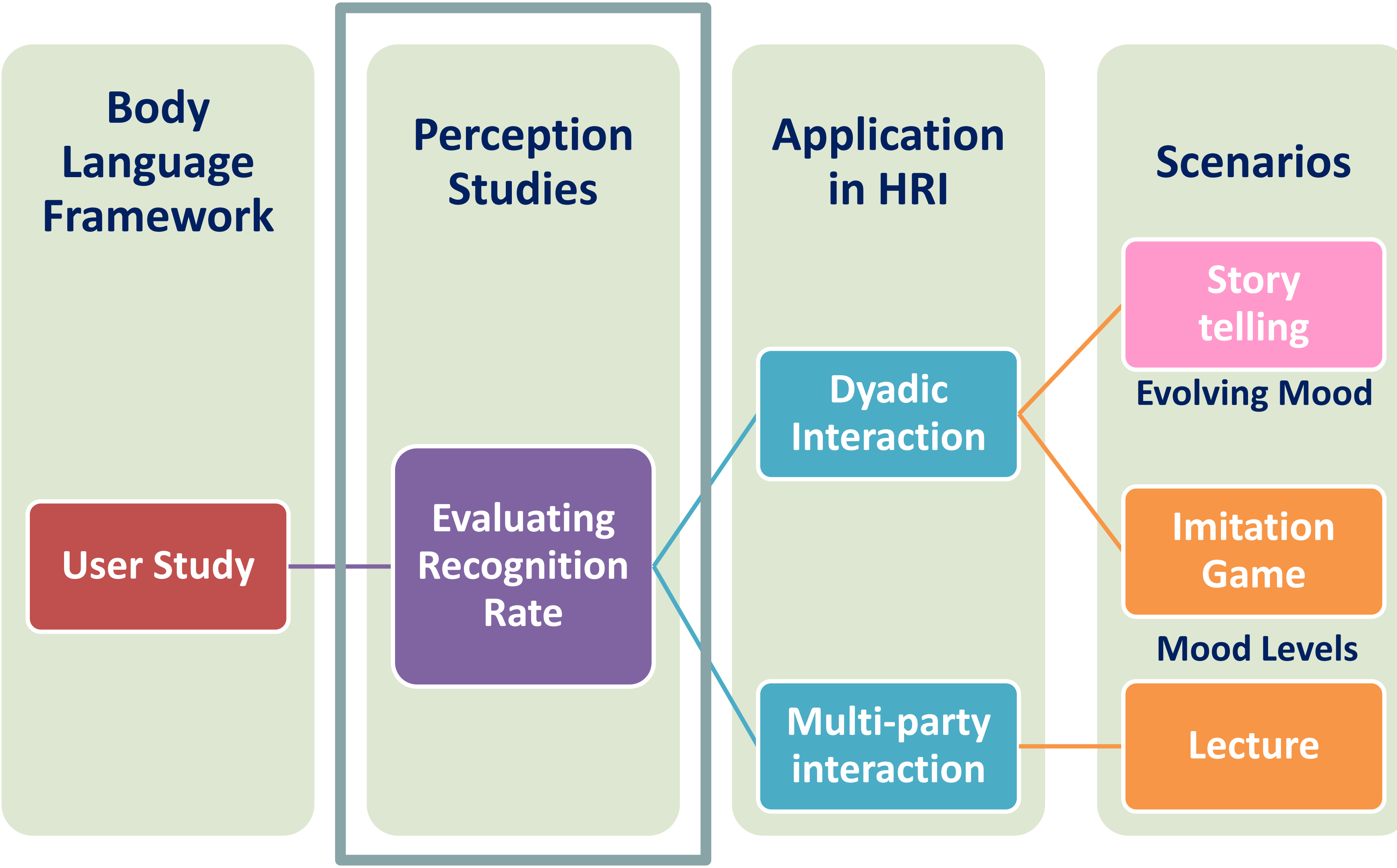
Definition of waving:
one hand swinging between two horizontally aligned positions repeatedly, where the palm should face forward.

Parameter	Negative	Positive
Hand-Height	low	high
Finger-Rigidness	bent	straight
Amplitude	small	large
Repetition	low	high
Hold-Time	long	short
Decay-Speed	slow	fast
Waving-Speed	slow	fast
Head-Up-Down	lower	raise
Head-Left-Right	look away	look at you

Modulated Waving Behavior



Overview



Body Language Framework
Perception Studies
Application in HRI
Scenarios
User Study
Evaluating Recognition Rate
Dyadic Interaction
Multi-party interaction
Story telling
Evolving Mood
Imitation Game
Mood Levels
Lecture

Recognition User Study

Study goals:

- *Recognition*: How do mood ratings (valence / arousal) of participants match with those expressed by modulated behaviors of the robot?
- *Parameter sensitivity*: How important are each of the modulation parameters for expressing mood?

Setup:

Subjects were shown robot body language unrelated to task or context

J Xu, J Broekens, K Hindriks, MA Neerincx, 2013. Bodily mood expression: Recognize moods from functional behaviors of humanoid robots. International conference on social robotics.

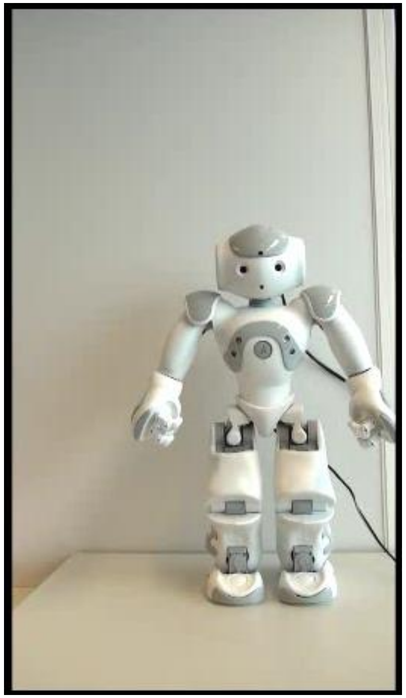
Hypotheses

- **H1:** People **can distinguish valence levels** from behaviors when **all behavior parameters (APS)** are modulated.

- **H2:** People can distinguish different valence levels when only **important parameters (IPS)** are modulated.
  People can still distinguish different valence levels when only modulating ***unimportant parameters (UPS)****, but the* ***recognition rate is worse*** *than the APS and IPS conditions.*

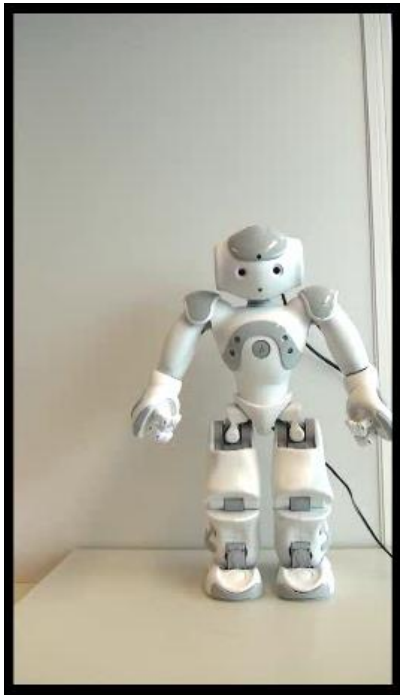
- **H3: Naturalness:** Behaviors generated by modulating all parameters (**APS**) are perceived as more **natural** than those generated by modulating only important parameters (**IPS**).

Waving – All parameters used

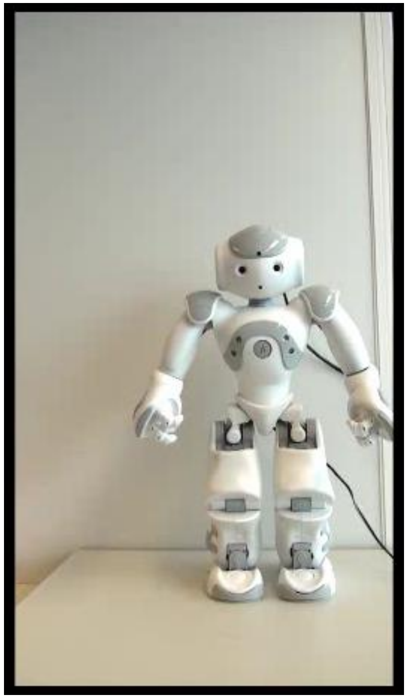


Very Happy

Neutral

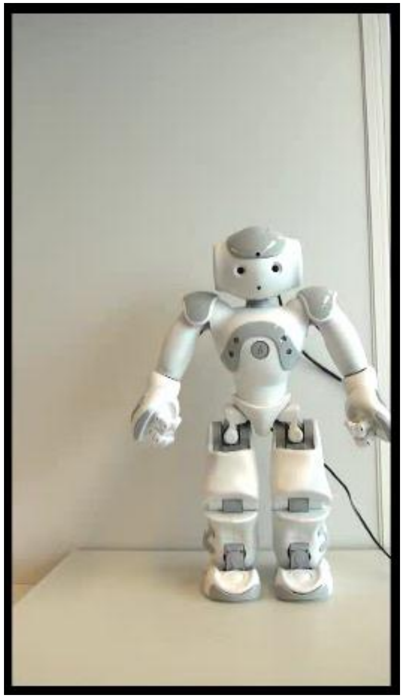


Very Unhappy (Sad)

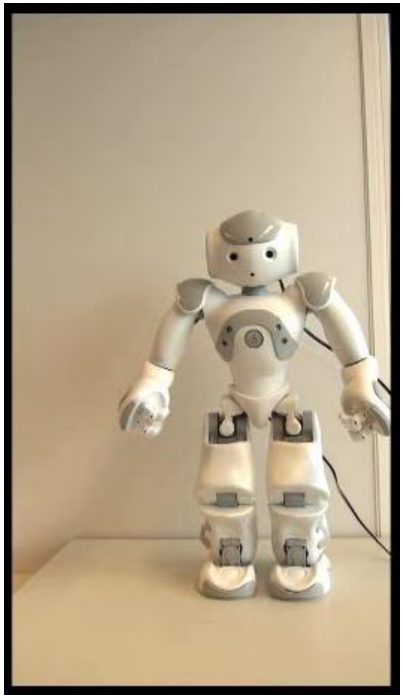


Angry

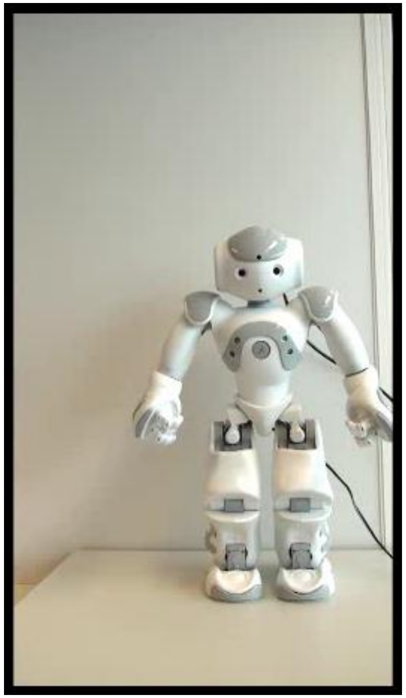
Pointing – All parameters used



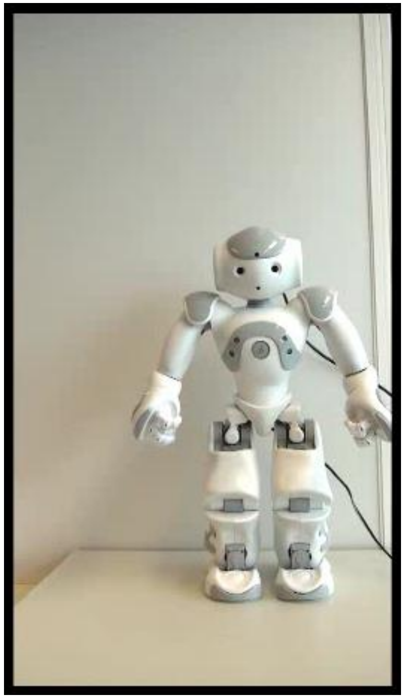
Very Happy



Neutral

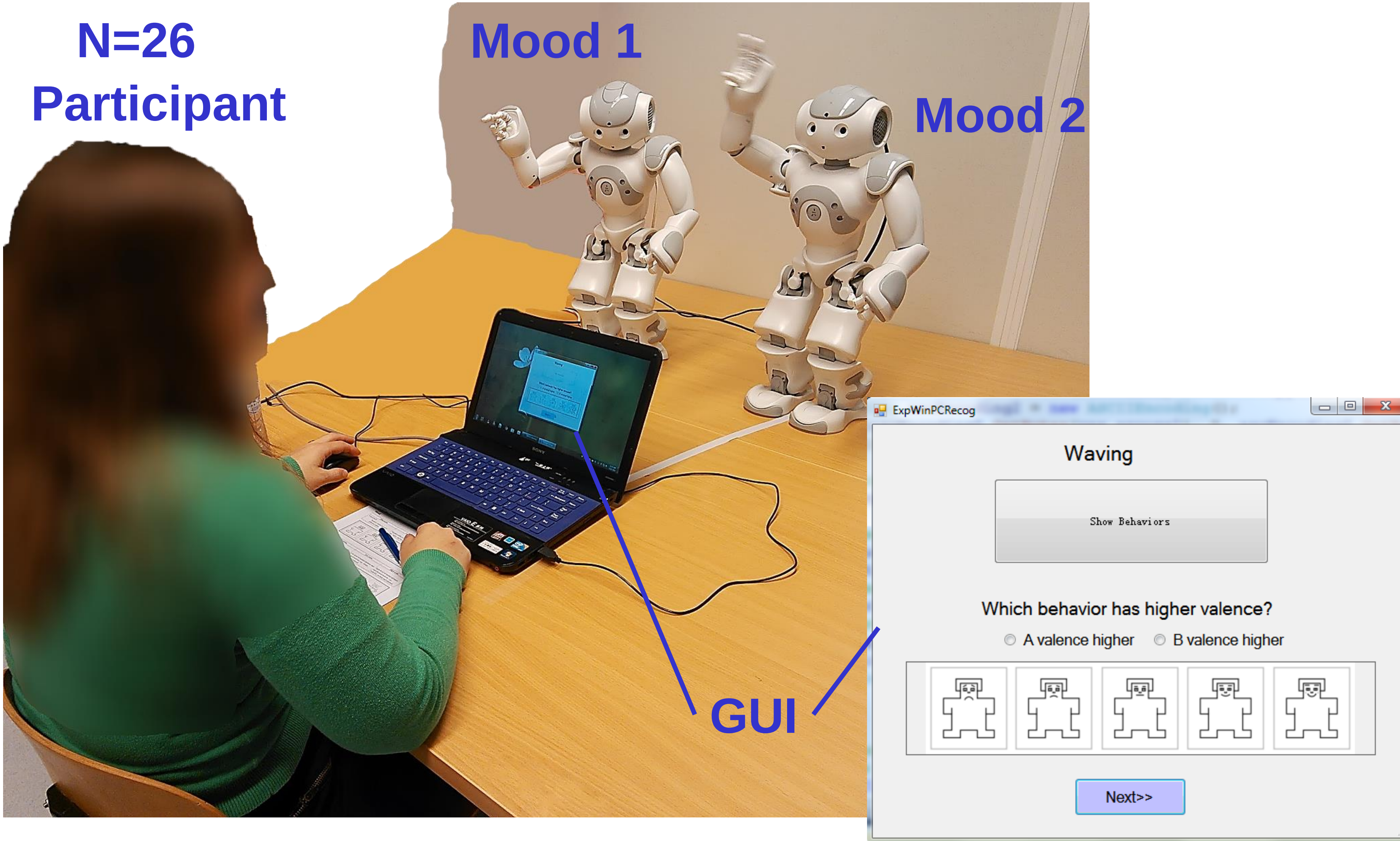


Very Unhappy (sad)

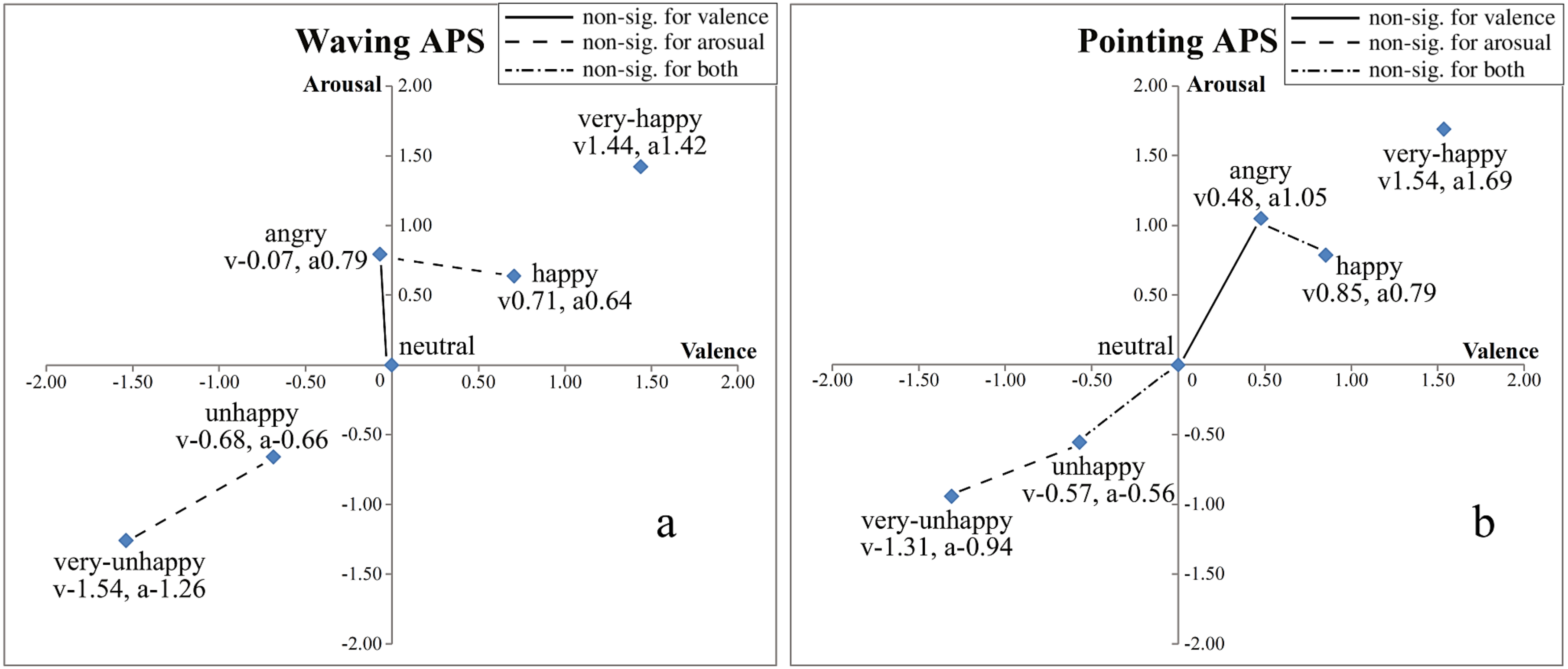


Angry

Experiment Setup for Paired Comparison



Paired Comparison Results

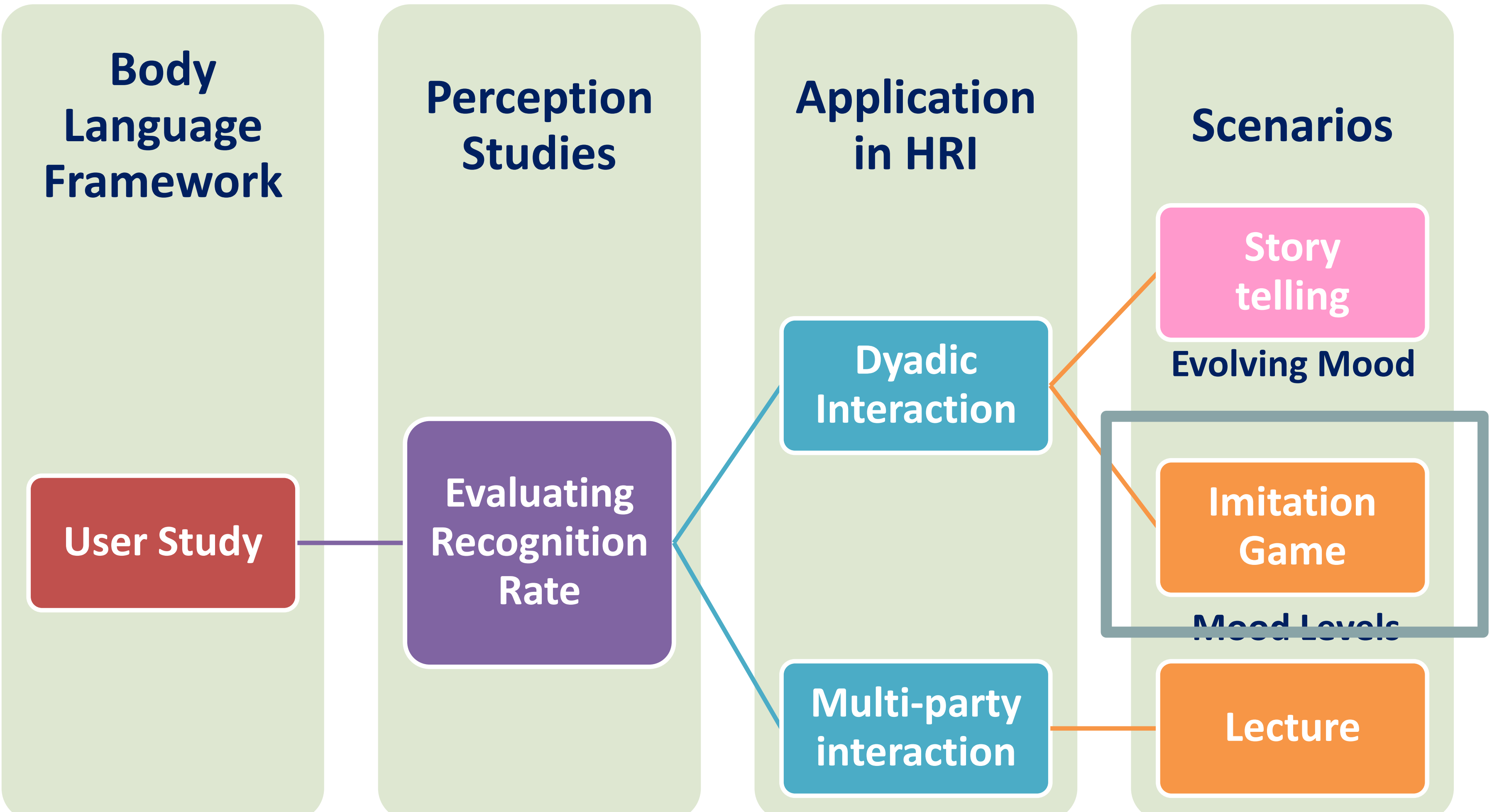


The valence / arousal of moods could be differentiated.

H1 and H2 are supported. H3 not.

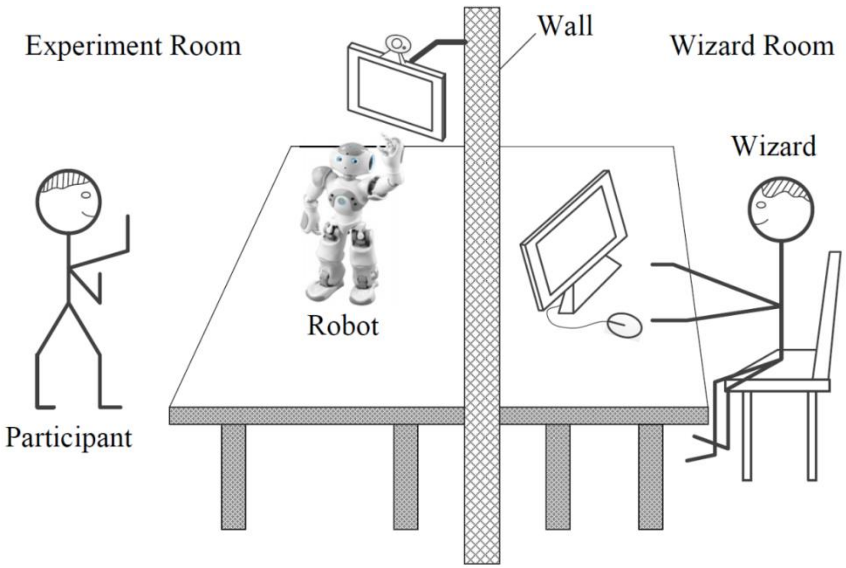
→no significant difference found between APS and IPS

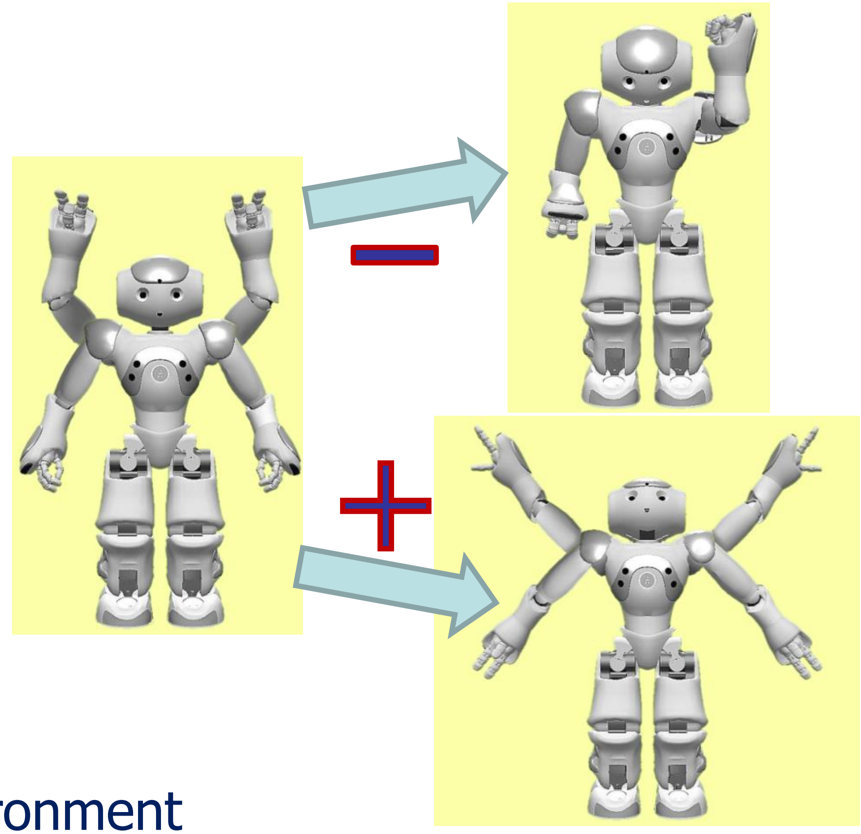
Overview



Body Language Framework
Perception Studies
Application in HRI
Scenarios
User Study
Evaluating Recognition Rate
Dyadic Interaction
Multi-party interaction
Story telling
Evolving Mood
Imitation Game
Mood Levels
Lecture

Evaluation in Context





- Functional behavior: game gestures
- Dyadic interaction in laboratory environment
- Recognition of robot mood
- Participants' mood: contagion effect
- Participants' game performance

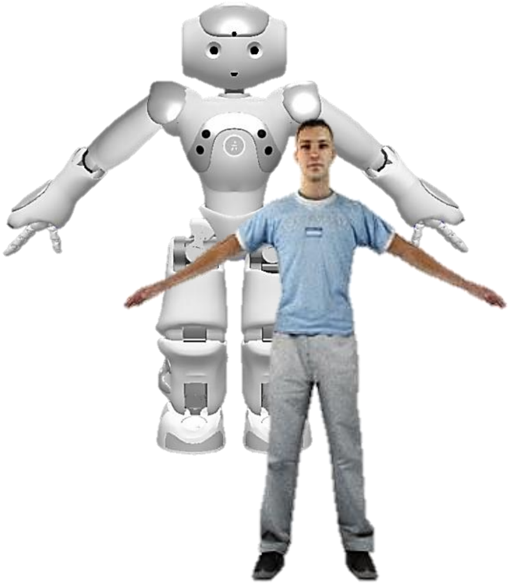
Xu, J., Broekens, J., Hindriks, K., Neerincx, M.A.: Robot Mood is Contagious: Effects of Robot Body Language in the Imitation Game. Autonomous Agents and Multi-agent Systems (AAMAS), 2014.

Motivation

- **Perception**: Can people recognize robot mood expressed in behaviors during an interaction, rather than in a pure recognition task?
- **Effects**: Can mood expressed by a robot influence people's affective state and performance during interaction?
- **Cognitive load**: how does cognitive load mediate the recognition and the effects?

Imitation Game

- **Imitate gestures**:
  - left-up,
  - right-up,
  - left-down,
  - right-down, and
  - four combinations.

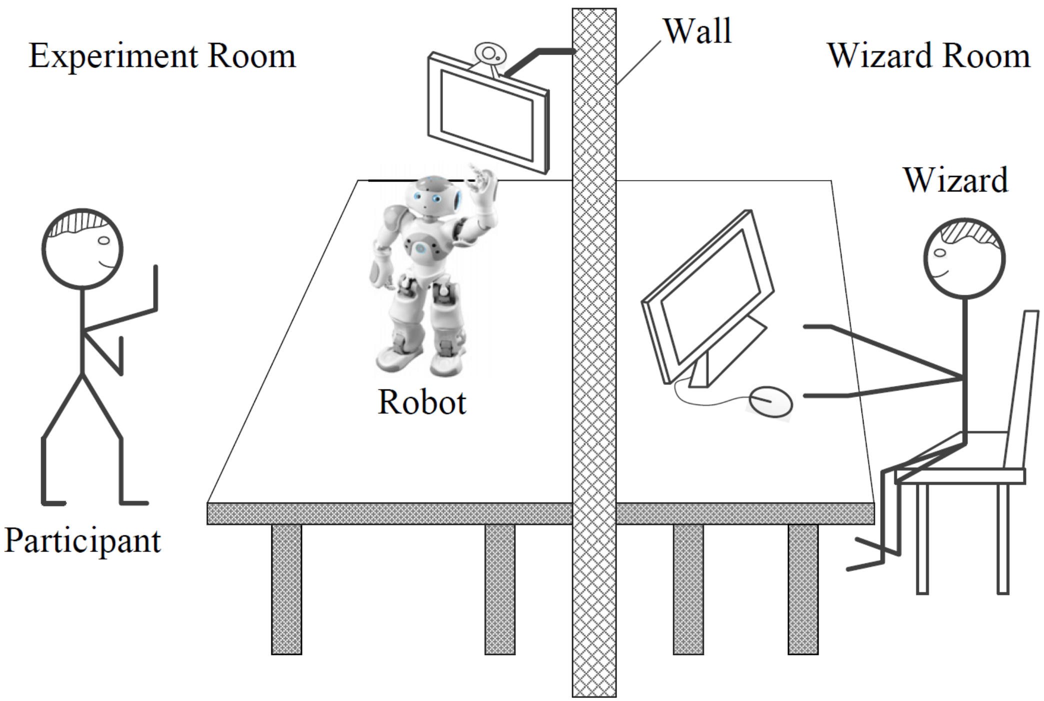


- A **memory** game: human players imitate gesture sequences (not exact movements) initiated by the robot, after the robot finishes a sequence.

- Varying game difficulty: sequence length & patterns.

Modulated Imitation Game Gestures

Wizard of Oz Setting



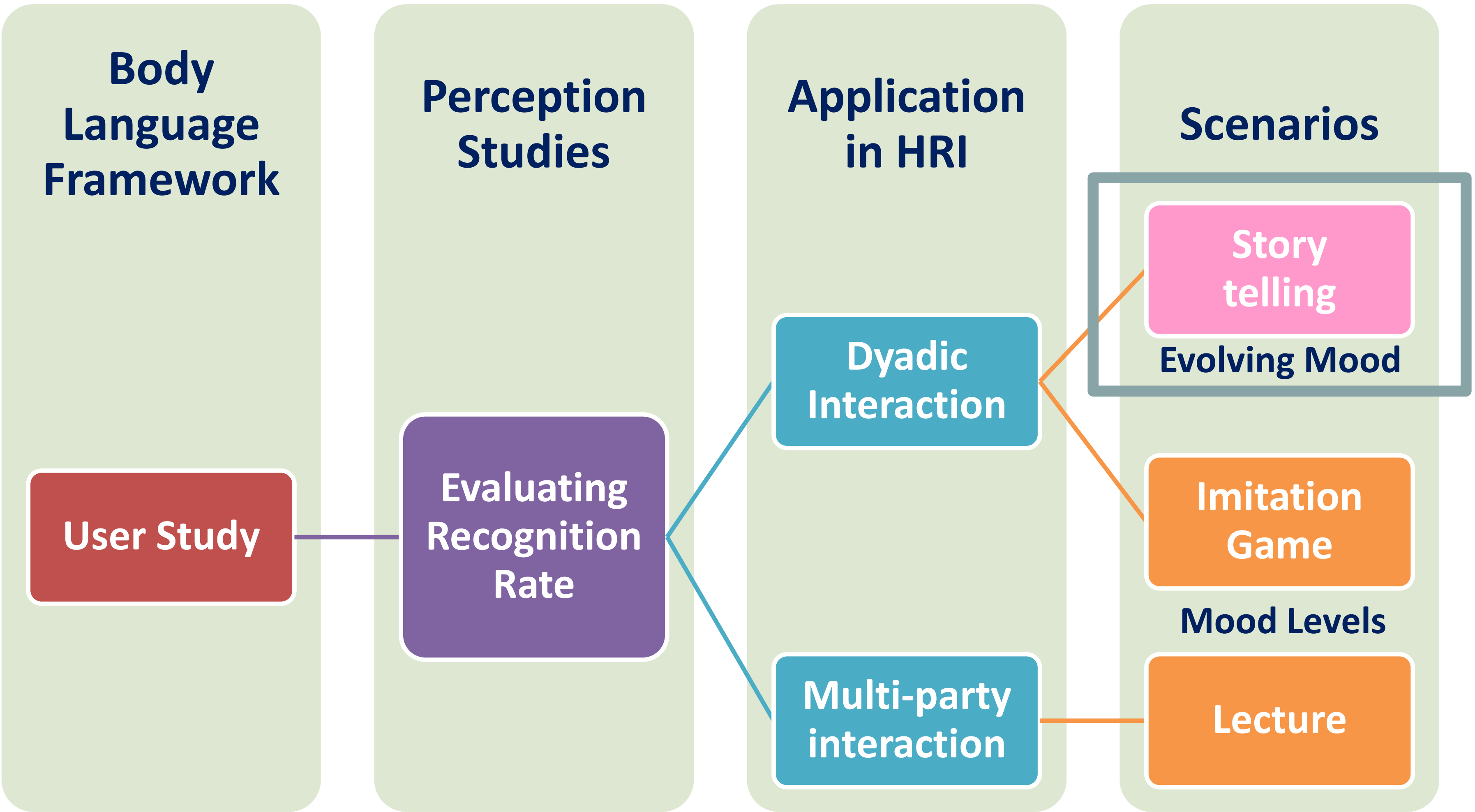
- Gestures were recognized by trained "wizard"
- Screen was used to guarantee S's gestures are in the view
  - Ease "robot" recognition
  - Excuse of the camera
- Neutral color: grey-white robot body; LED lights off

- 36 participants: 18 for each difficulty condition

Findings

- We were able to use parameterized behavior to express robot mood in interaction scenario.
- Model generalized to functional behaviors (gestures).
- People are able to distinguish between positive and negative robot mood.
- Mood contagion takes place between the robot and the human, evidenced by self-reported mood and game performance.

Overview


Body Language Framework
Perception Studies
Application in HRI
Scenarios
User Study
Evaluating Recognition Rate
Dyadic Interaction
Multi-party interaction
Story telling
Evolving Mood
Imitation Game
Mood Levels
Lecture

Motivation

- Express mood that is evolving over time.

- In a multimodal communication context.
  - Semantics
  - Body language
  - Voice (not manipulated)

- Improve storytelling experience.

Hypotheses

- **H1**: Listeners can distinguish whether the robot body language expresses mood that follows the story mood (consistency ratings).
- **H2**: Consistent expression enhances perception of the story mood.
- **H3**: Consistent expression strengthens mood induction.
- **H4**: Consistent expression improves story telling experience.

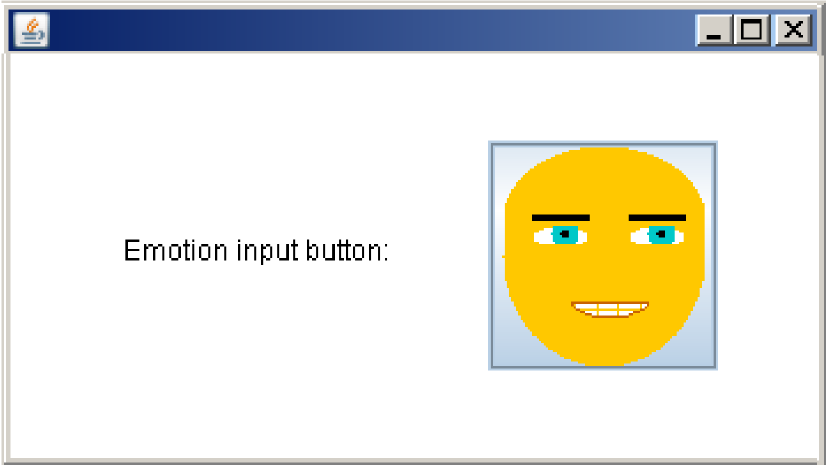
Experimental Design

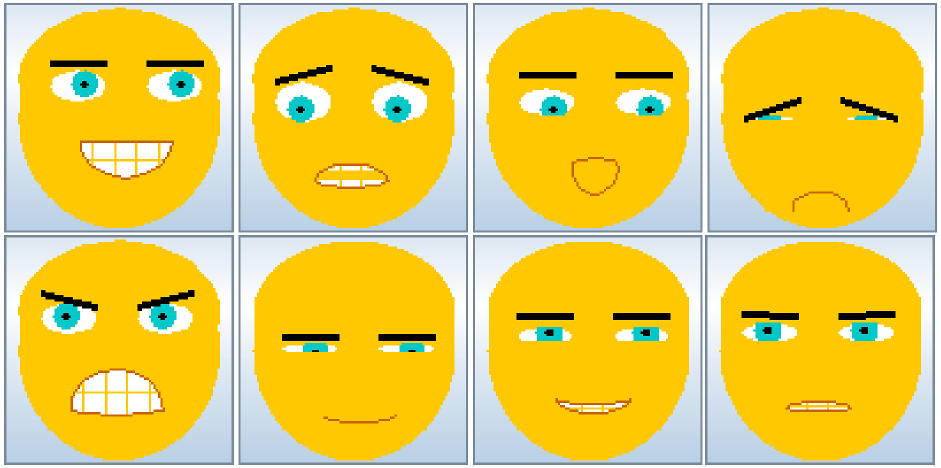
- Each subject listened to two stories in counter-balanced order
  - one realistic story (1 min 45 sec)
  - one fantasy story (3 min)
- 66 participants (42 males; 24 females)
- Age range from 19 to 48 (Mean = 28.0, SD = 4.8)

Between-subject design

Conditions	Expression
Control	No gestures; only (random) leg movements
Consistent	Gestures express mood that is **consistent** with story mood
Inconsistent	Gestures express mood that is **inconsistent** with story mood

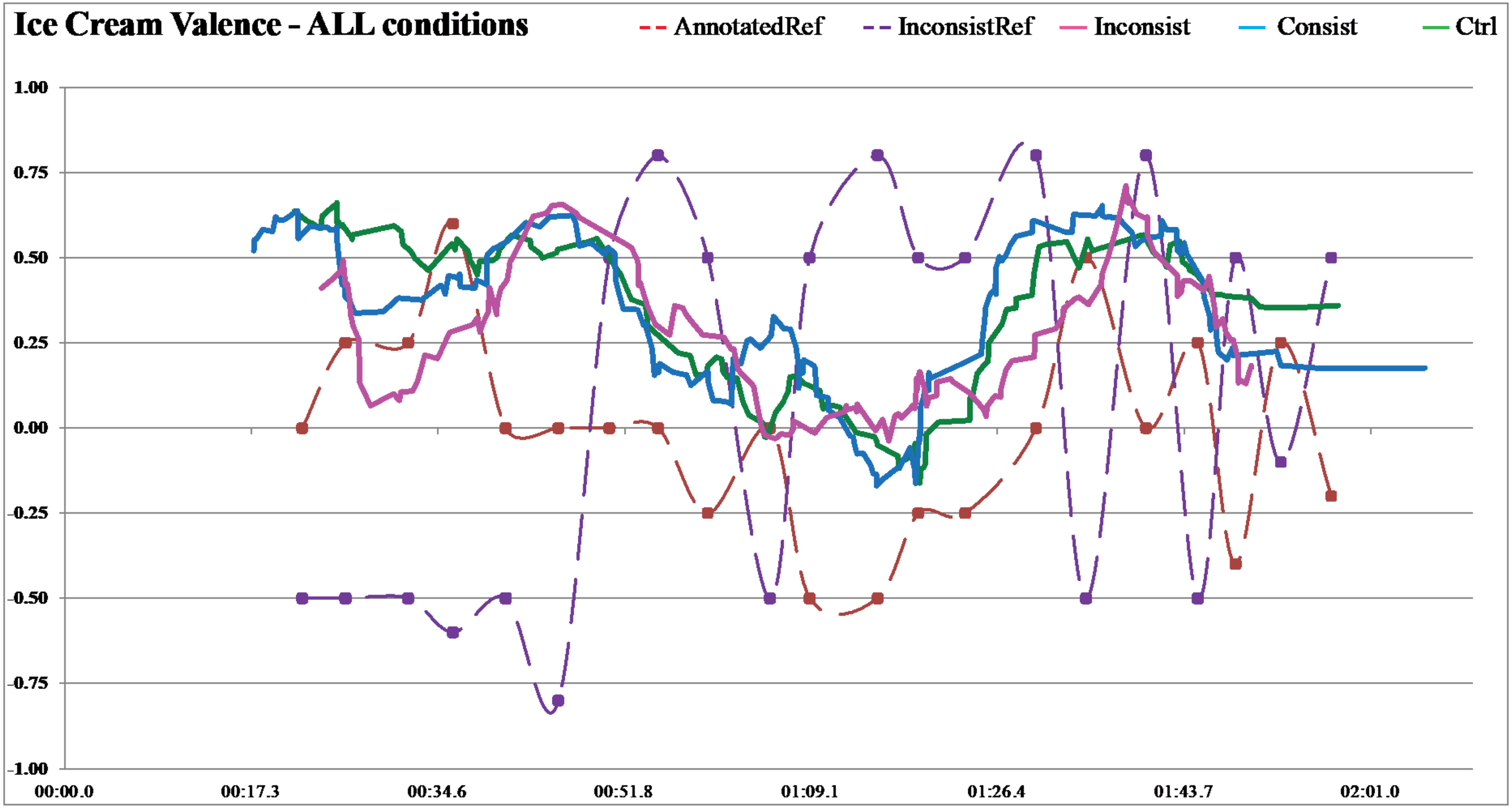
Track Story Mood using AffectButton





- Track the mood of the story in real time using AffectButton: input whenever think the story mood is changing
- A training session was provided

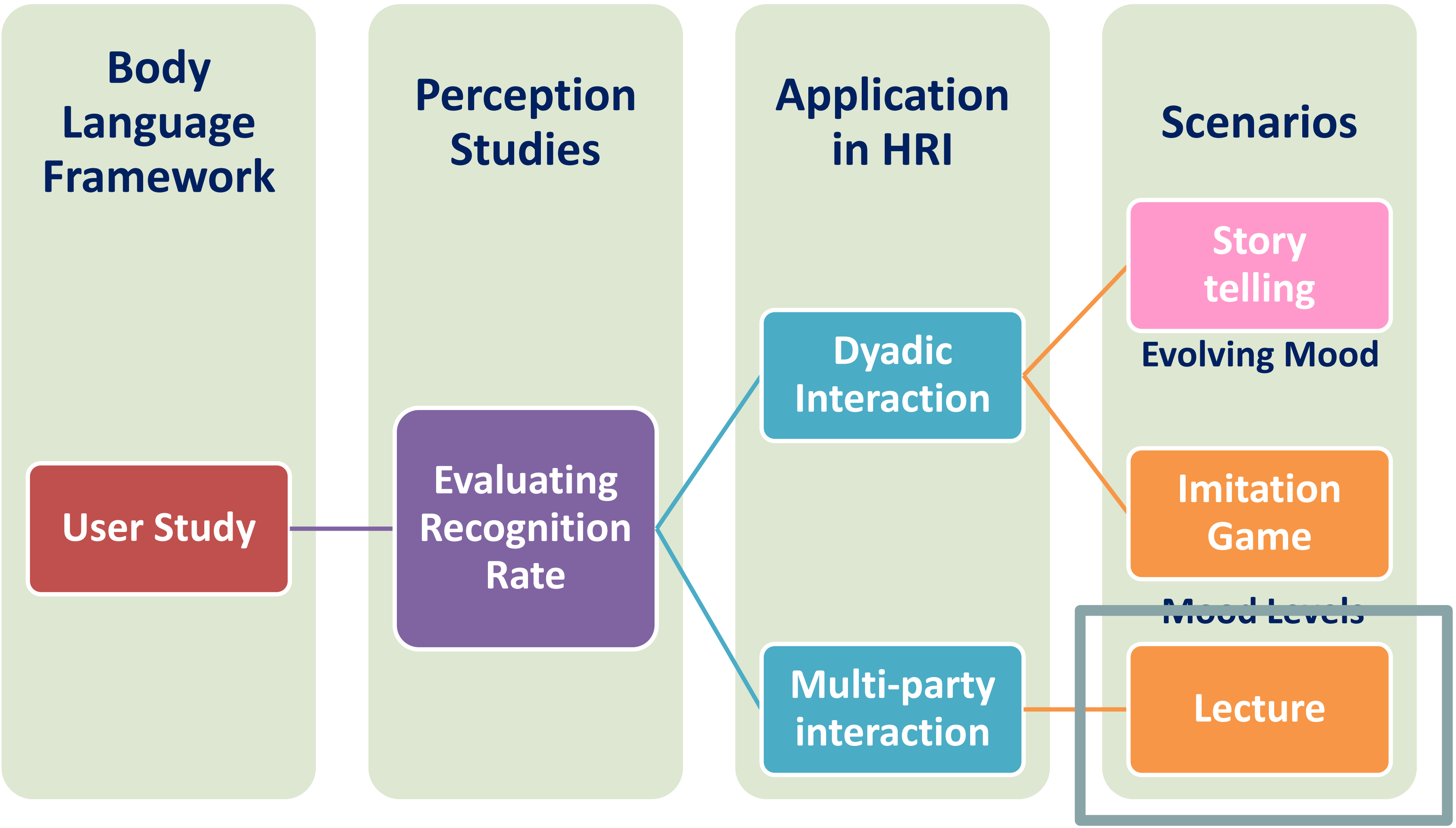
Trace of Perceived Story Mood



Findings

- Affective body language is able to express mood that is evolving over time
- Consistent body language helps capture story mood and makes the mood stronger
- Semantic channel appears to take priority over body language: some participants ignored the body language when they found it incongruent
- Affective body language influences mood induction process
- Consistent body language improves the storytelling experience

Overview



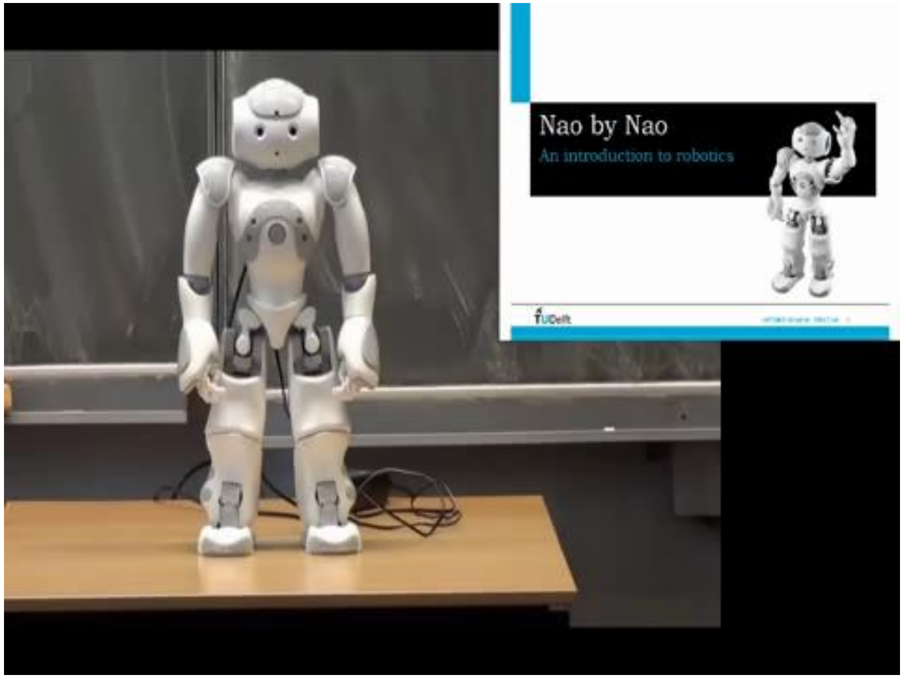
Body Language Framework
Perception Studies
Application in HRI
Scenarios
User Study
Evaluating Recognition Rate
Dyadic Interaction
Multi-party interaction
Story telling
Evolving Mood
Imitation Game
Mood Levels
Lecture

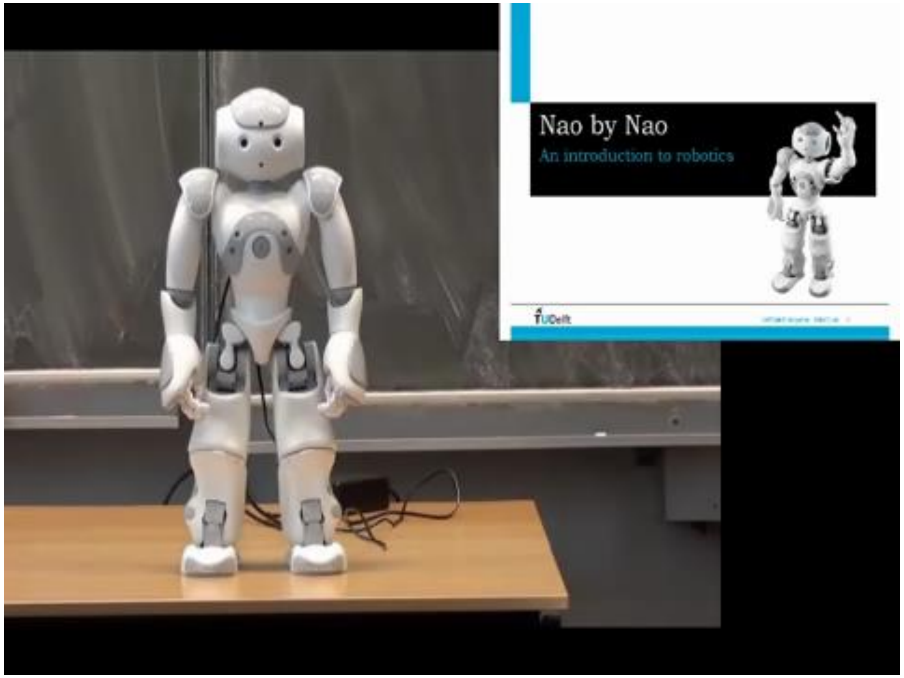
Motivation

- Expressing mood continuously over time
  - Expressing mood across a broad range of functional behaviors (41 gestures)
  - Expressing mood over a longer period (30 min)

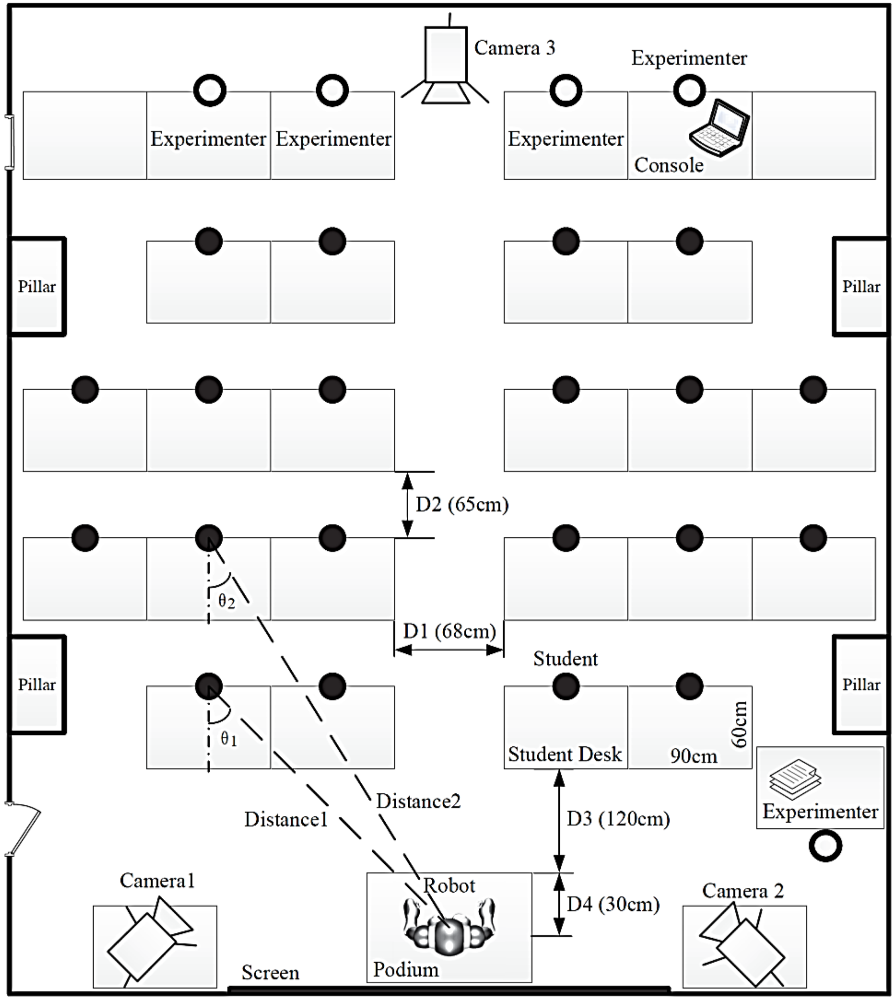
- A real-life setting (“in the wild”):
  - One-on-multiple interaction: group effects
  - Participants were not primed to pay attention to any form of affective expression
  - Between-subject design
  - Uncontrolled factors

Happy or Sad


Nao by Nao
An introduction to robotics


Nao by Nao
An introduction to robotics

Experimental Setup

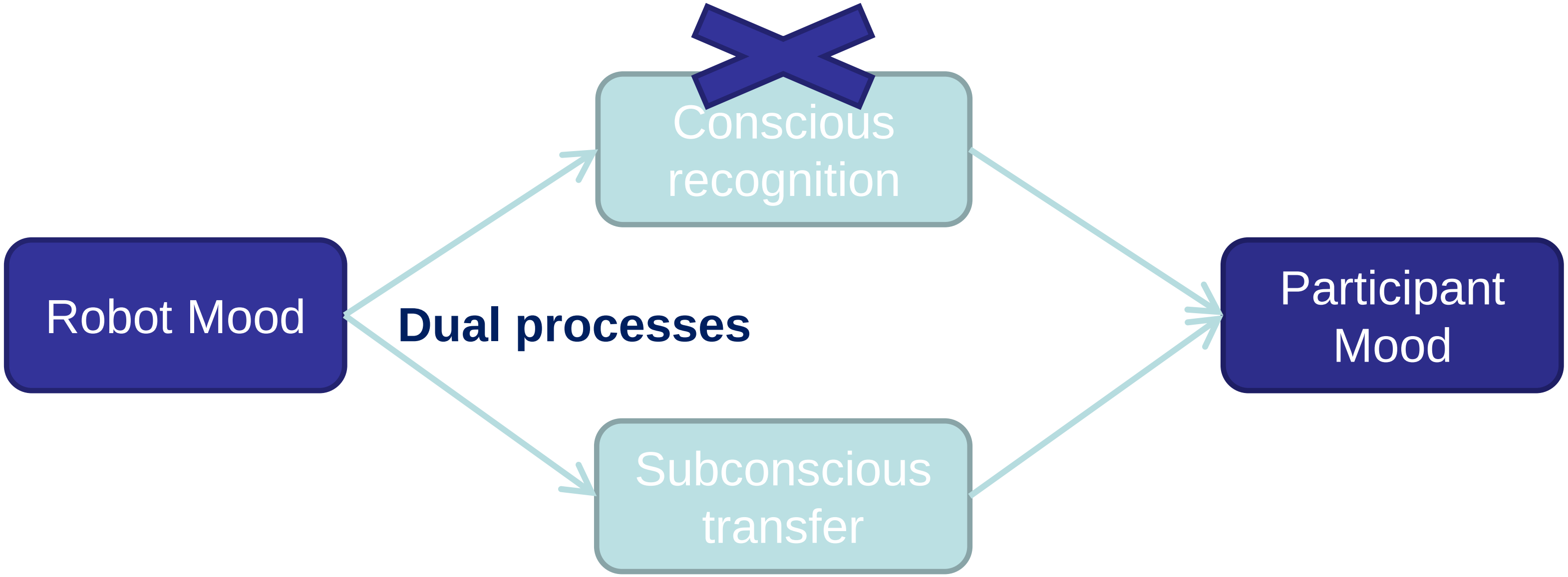


- Small lecture room: 26 seats
  - Close to robot
  - Usual classroom
- Seat layout in grid pattern
  - Distance
  - Angle
- Video recording
- Course materials

Findings – Recognition

- Did not observe a clear difference in perceived robot mood
- Attention paid to lecture slides/listening
- Gesture modulation attributed to teaching quality
- Robot mood attributed to various non-manipulated factors (voice, etc.)
- Small sample size for between-subject design

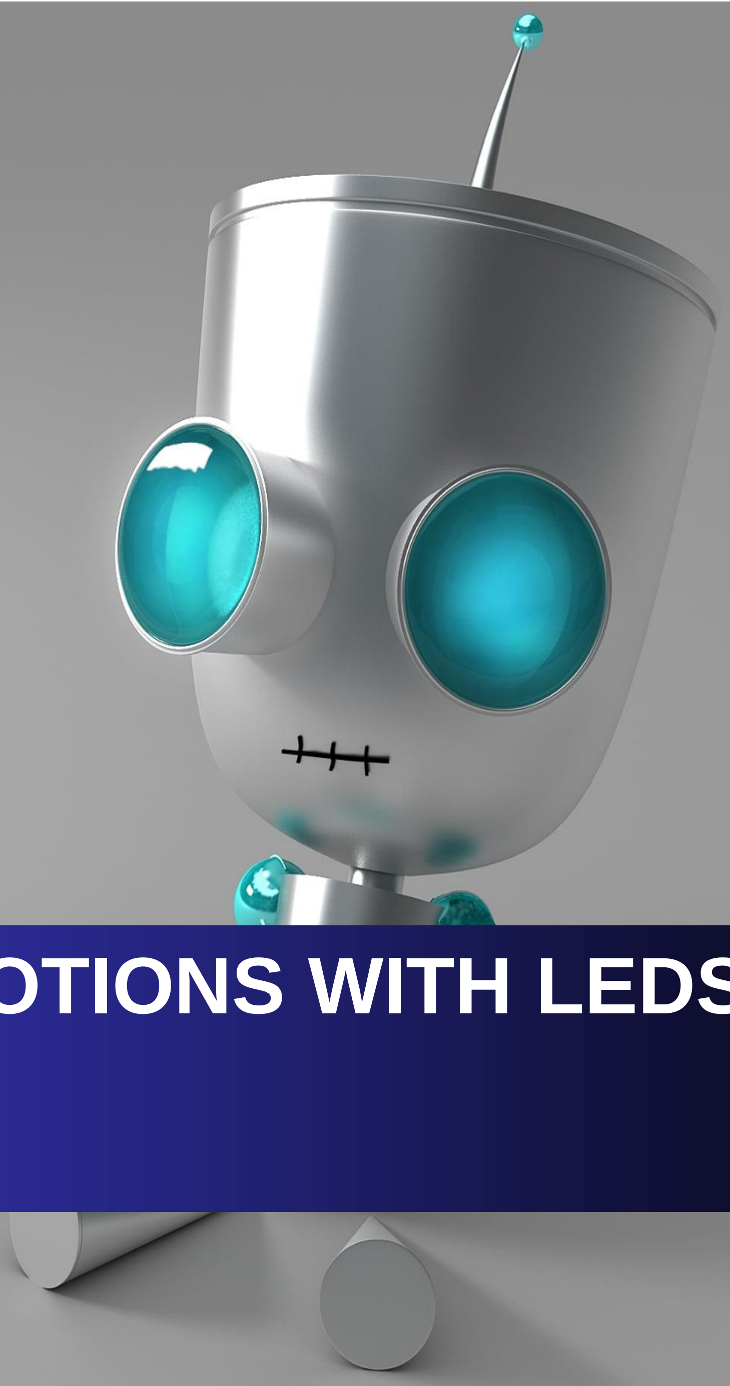
Findings – Contagion



Findings – Performance

- Six multiple-choice quiz questions about the lecture content

- We **did not observe an effect** of robot mood on students’ task performance



EXPRESSING EMOTIONS WITH LEDS

Motivation

Minimal generic and parametric model for emotion expression

- Minimal → as few parameters as possible
- Generic → robot-independent
- Parametric → function of valence & arousal, covers entire valence-arousal space

Minimal Set of Parameters

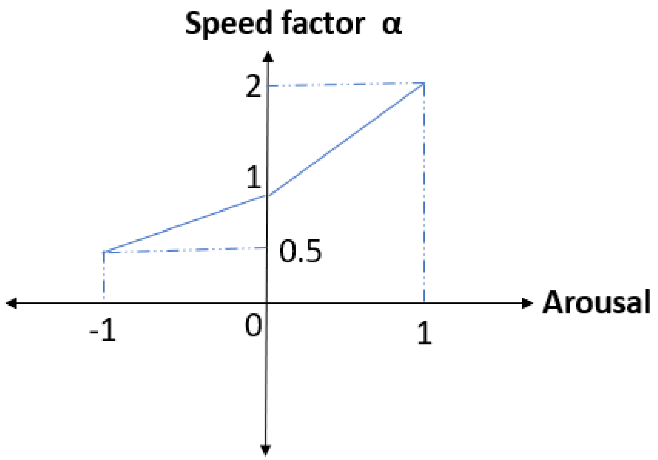
- Motion features:
  - speed,
  - amplitude,
  - repetition
- Body language features:
  - vertical head pose
  - bend/straight

Feature Specification

- Representation
  - Pose – angle configurations at an instant $\boldsymbol{\theta}_i$
  - Gesture – time series of poses

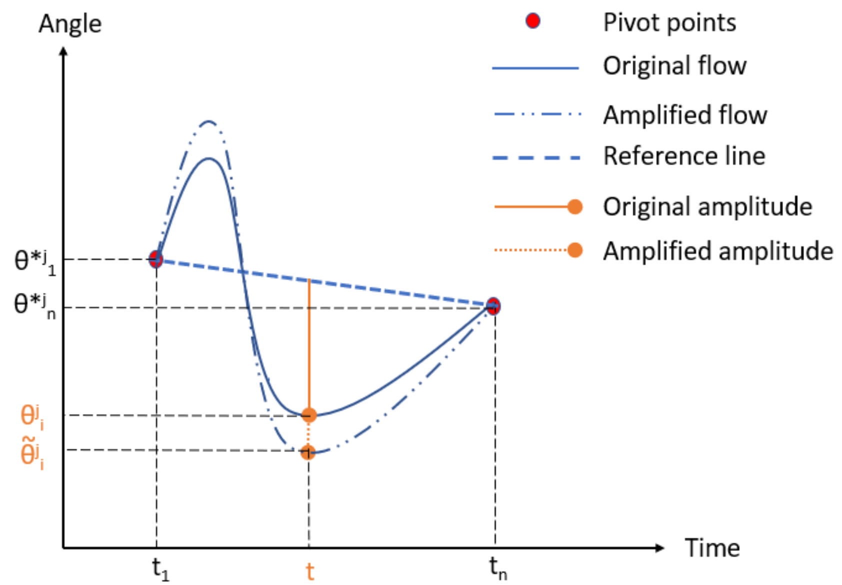
    $$\text{poses} = \langle\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n\rangle \quad t = \langle t_1, \ldots, t_n\rangle$$
  - Affect – a point in valence-arousal space

- Speed
  - Modulated as
  - $\alpha > 1 \rightarrow$ faster, $\alpha < 1 \rightarrow$ slower
  - Arousal oriented feature

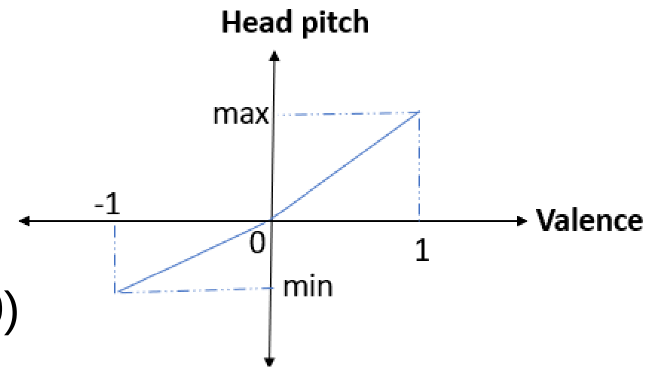


Feature Specification

- Amplitude
  - α > 1 → amplified, α < 1 → reduced
  - Valence oriented feature



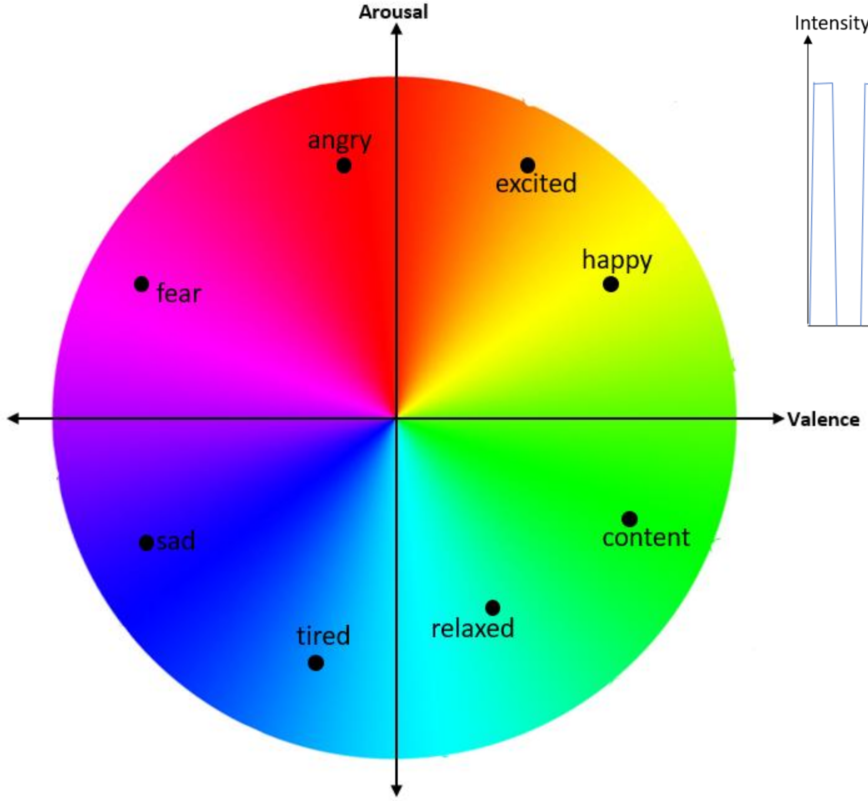
- Vertical head pose
  - Modulates head pitch
  - Only affects in Q1(valence & arousal > 0) and Q3 (valence & arousal < 0)

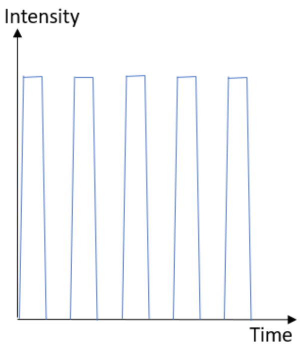


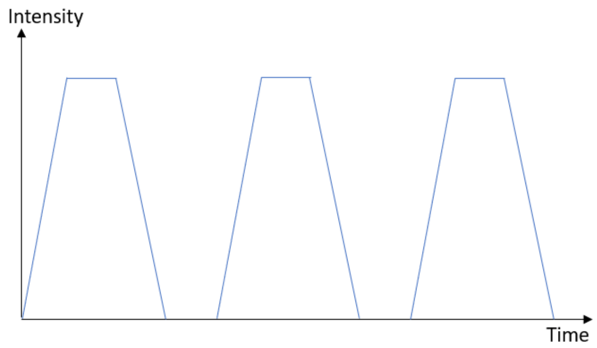
LEDs

LED patterns

- Colour – depends on both valence & arousal
- Blink frequency, waveform – arousal oriented

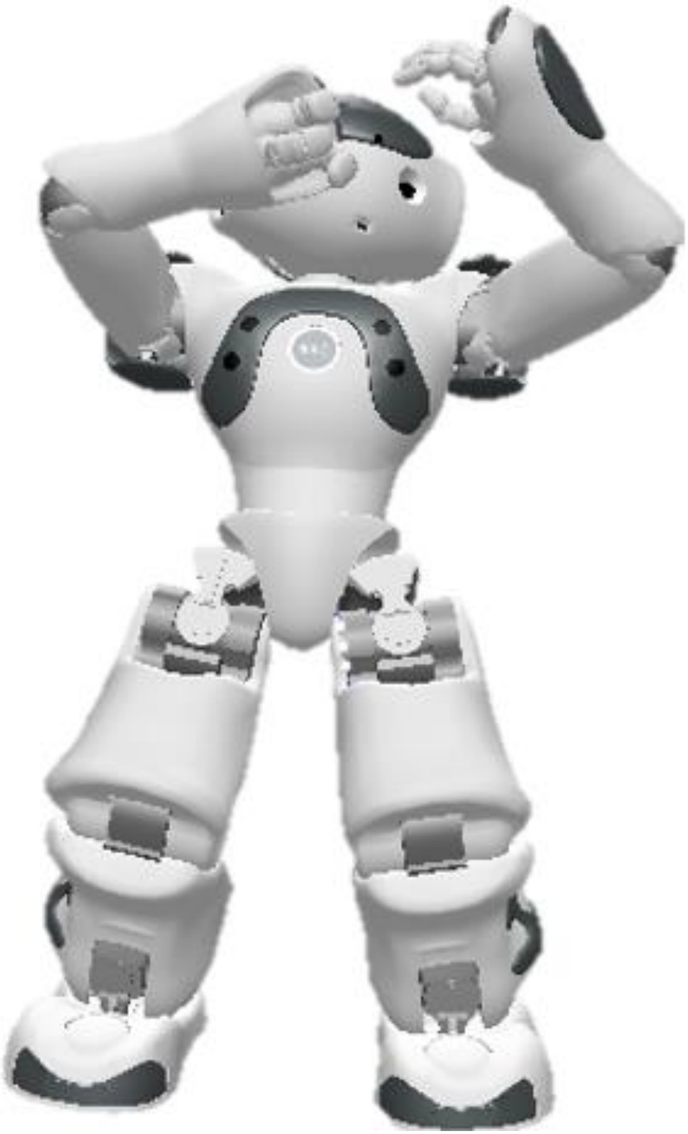




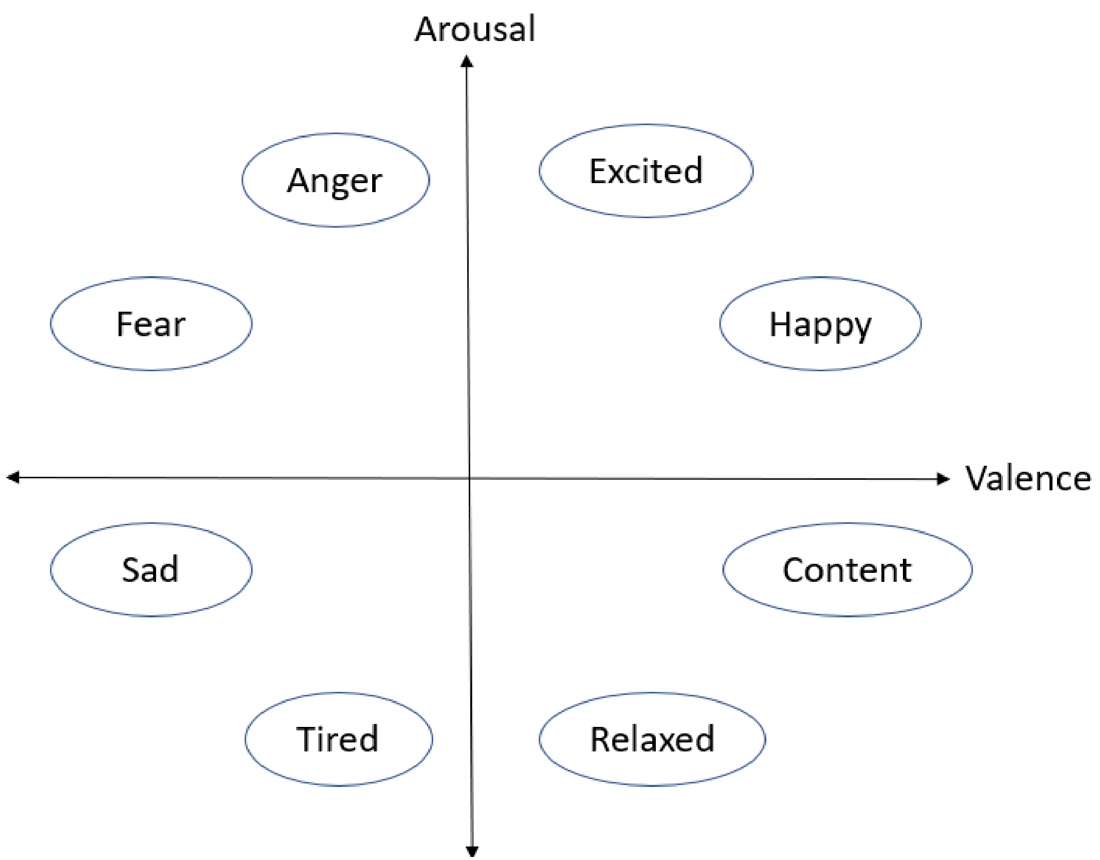


Fagerberg et al. *emoto: emotionally engaging interaction*, 2004.
Terada et al. *Artificial emotion expression for a robot by dynamic color change*, 2012.

Expressing Fear



Evaluation using Mturk



Affect list

- Only for testing
- 8 affects – excited, happy, content, relaxed, tired, sad, fear, angry

Setup

Type	Iconic	Metaphoric	Deictic	Beats
Gesture	Wave, Handshake, Look-around	Nod-yes, Clap	Pointing	These, This-or-that

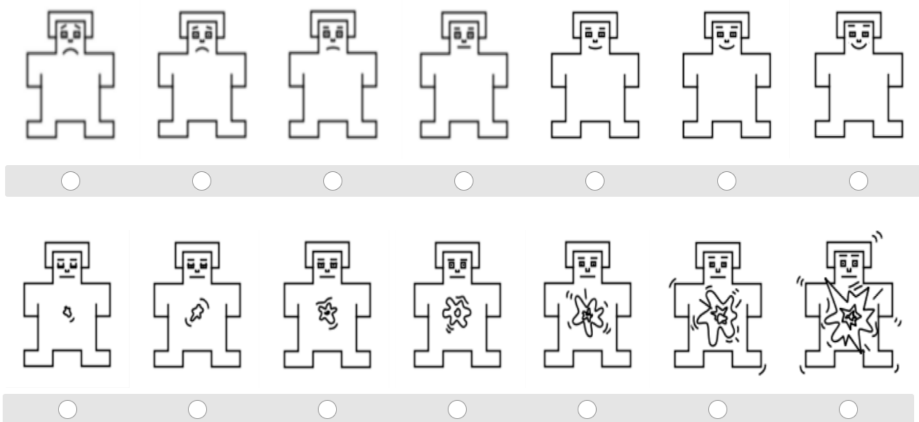
Phase	1	2	3
Features	Motion, Body language	+ LED patterns	Pose repertoires
Affects	All	Not recognized in phase 1	Neither recognized in phase 1 nor 2

David McNeill. *Hand and mind: What gestures reveal about thought*, 1992

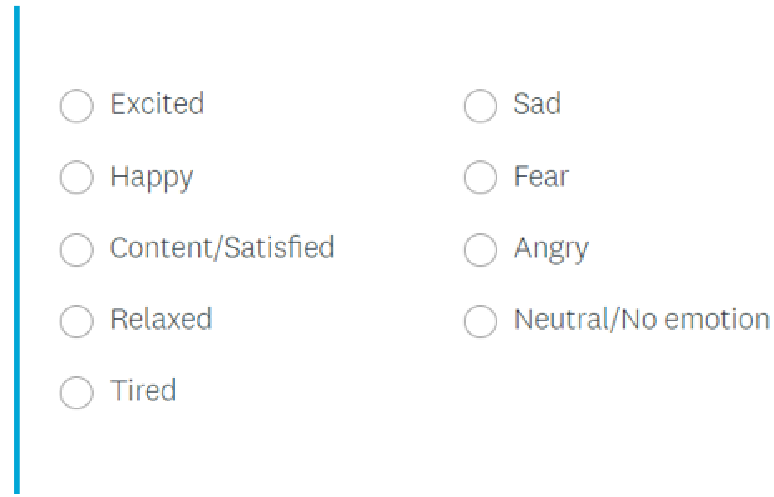
Materials & Measures

Valence, Arousal ratings

SAM questionnaire

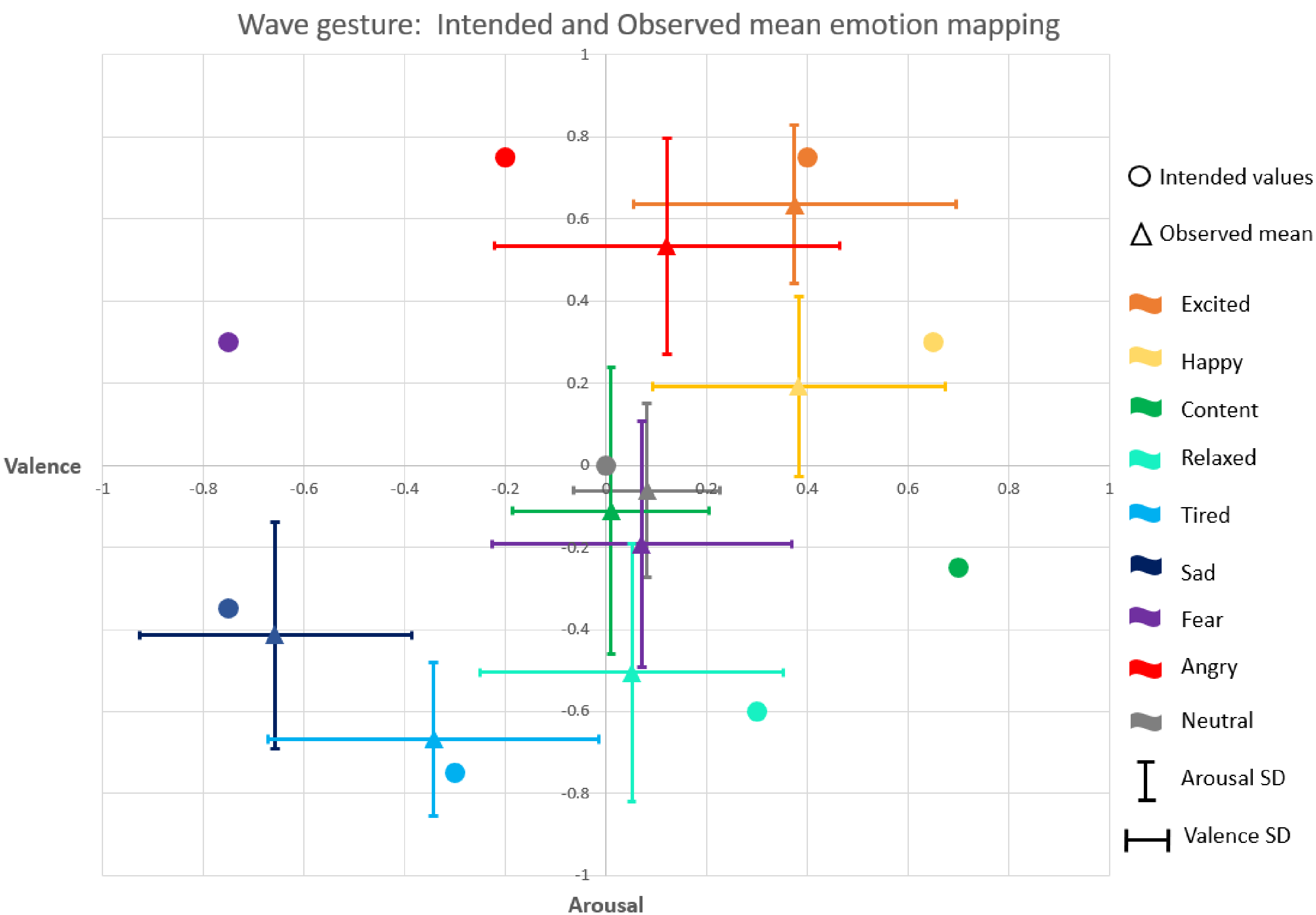


Affect labels – Forced choice



- Survey on Amazon Mechanical Turk
- 8.5 seconds long videos of Nao

Analysis: Waving



Intended affect	# recognised
Neutral	25
Happy	17
Excited	29
Anger	3
Fear	0
Sad	28
Tired	19
Relaxed	9
Content	6

Summary of Findings

- 7 out of 8 affects perceived
  - Motion only: excited, happy, sad, tired
  - +LEDS: angry, relaxed
  - Key pose: fear
  - content: high valence, low arousal, is very difficult to express

- Motion & body language features – strong indicators of arousal, valence needs more features

- LED patterns – strong indicators of arousal, improves valence ratings

Summary

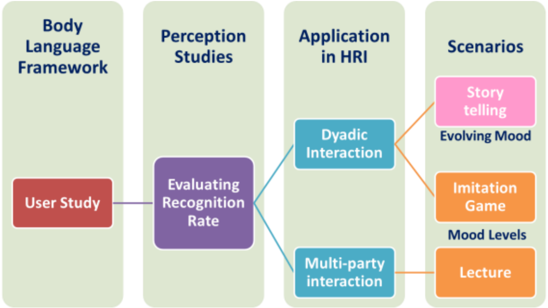
- NVC is essential for *embodied interaction* with humans
  Contributes to fluidity, transparency, trust,…

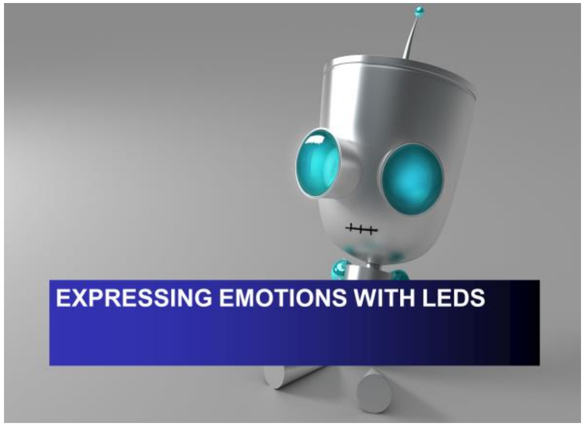
- NVC modalities include (but are not limited to): gaze (mutual gaze, deictic, joint attention, …), gestures (iconic, metaphoric, beat, …), expressive motion, proxemics (interpersonal distances and personal space), haptics (physical HRI), and prosody

- Robot-specific modalities include lights, sound, and robot-specific motion

- NVC is still a vast open area of research in HRI

Summary





Literature:

- Shane Saunderson, Goldie Nejat (2019) How Robots Influence Humans: A Survey of Nonverbal Communication in Social Human–Robot Interaction