Simulating affective processes for autonomous agents: the WASABI[1] architecture

Dr. Christian Becker-Asano

(formerly University of Bielefeld, Germany)

Intelligent Robotics and Communication Labs
Advanced Telecommunications Research Institute
Kyoto, Japan

[1][W]ASABI: [A]ffect [S]imulation for [A]gents with [B]elievable [I]nteracivity

Overview

I. Introduction

- Motivation
- Terminology
- The embodied agent MAX

II. Interdisciplinary background

- Psychological background
- Neurobiological background

III. Affect Simulation

- Emotion dynamics for MAX as a museum guide
- Primary emotions in the Skip-Bo scenario
- Integrating secondary emotions → WASABI
- Empirical evidence on the effect of secondary emotions

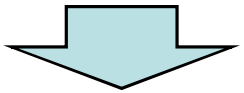
IV. Summary and ongoing work

Motivation

Why to simulate affect for embodied agents:

1. The "Believable-agent-motive":
   "Embodied conversational agents that show emotions in the way they act or behave in environments where they interact with humans [..] are more believable and engaging than similar agents that do not show emotions."
2. The "Experimental-theoretical-motive":
   "The system is built and used as an experimental environment to verify or falsify hypotheses based on the theoretical insights expressed in the emotion theory."

(Burghouts, op den Akker, Heylen, Poel & Nijholt 2003)



The embodied conversational agent MAX:

1. as an interactive & believable embodied agent
2. as testbed for emotion theory

Terminology: What is "affect simulation"?

The term affect subsumes (see also PART II):

1. Emotions (short duration)
2. Mood (longer duration)
3. Personality traits (nearly stable dispositions)

What is "Affective Computing"?

"Computing that relates to, arises from, or deliberately influences emotions" (Picard 1997, p. 13)

The embodied conversational agent MAX:

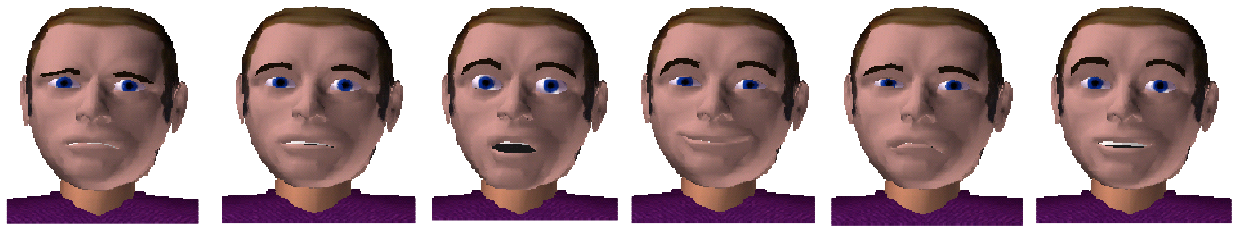
1. as an interactive & believable embodied agent
2. as testbed for emotion theory

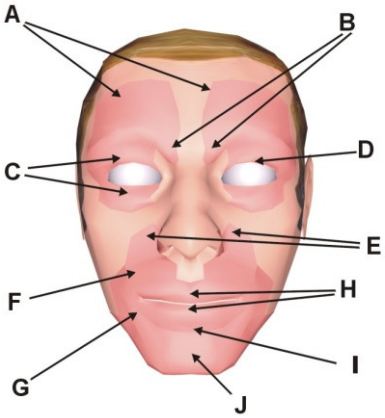
Embodied agent MAX (Bielefeld, Germany)

PART II

Kinematic skeleton with 53 degrees of freedom

- Model-based body animation (Kopp 2003)
- Cognitive architecture based on a Belief-Desire-Intention (BDI) framework (Leßmann 2002)
- Text-to-speech system realizes lip-sync facial animation of expressive speech (Stößel 2001)
- Facial animation:
  - based on 21 simulated muscles in accordance with „Action Units" (Ekman/Friesen)
  - Six muscle sets defined for emotional expressions:





Interdisciplinary background

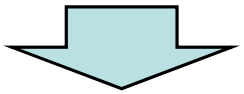
1. Psychological background
   - "OCC-model" of emotions (Ortony, Clore & Collins 1988)
   - Component Process Model (Scherer 1984/2006)
   - "Core Affect" (Russel & Feldmann Barrett 1999)
     & Pleasure-Arousal-Dominance space (Russel & Mehrabian 1977)
2. Neurobiological background (Damasio 1994)
   - Primary emotions
   - Secondary emotions

The OCC-model of emotions

A structural theory of emotions

Necessary conditions for the emergence of emotions:

1. Situations that cause the emotions
2. Persons that experience the emotions
3. The appraisal of the situation by the person

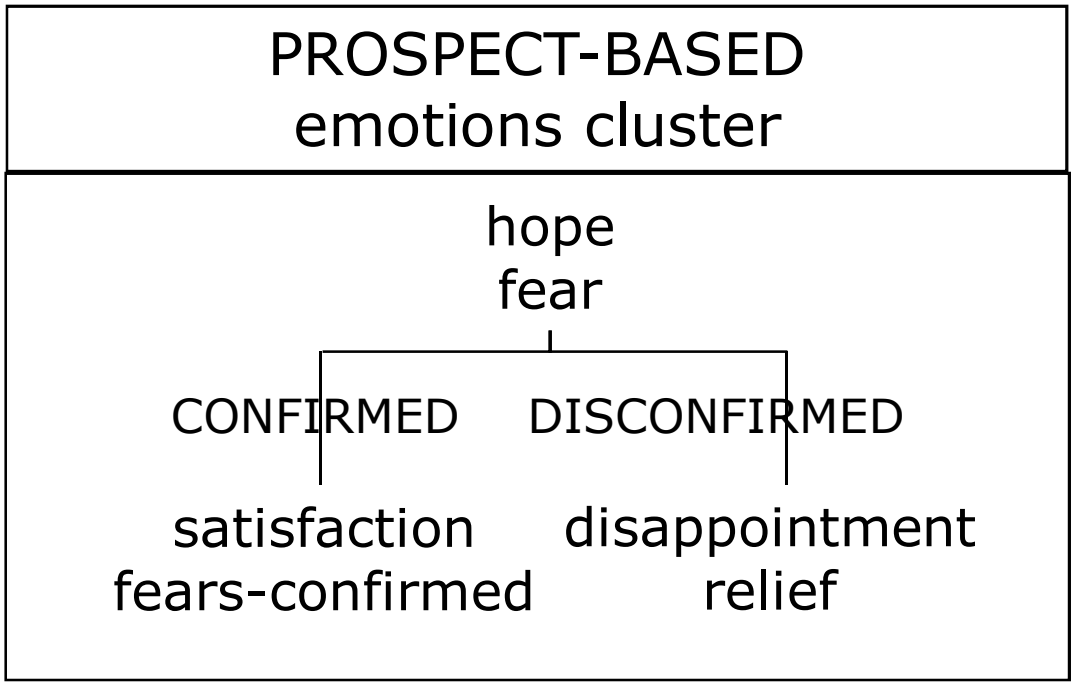


Appraisal theory
proposing 22 distinct emotions

(Ortony, Clore & Collins 1988)

OCC-model: Six prospect-based emotions

IF event-has-consequences-for-self AND prospects-relevant THEN



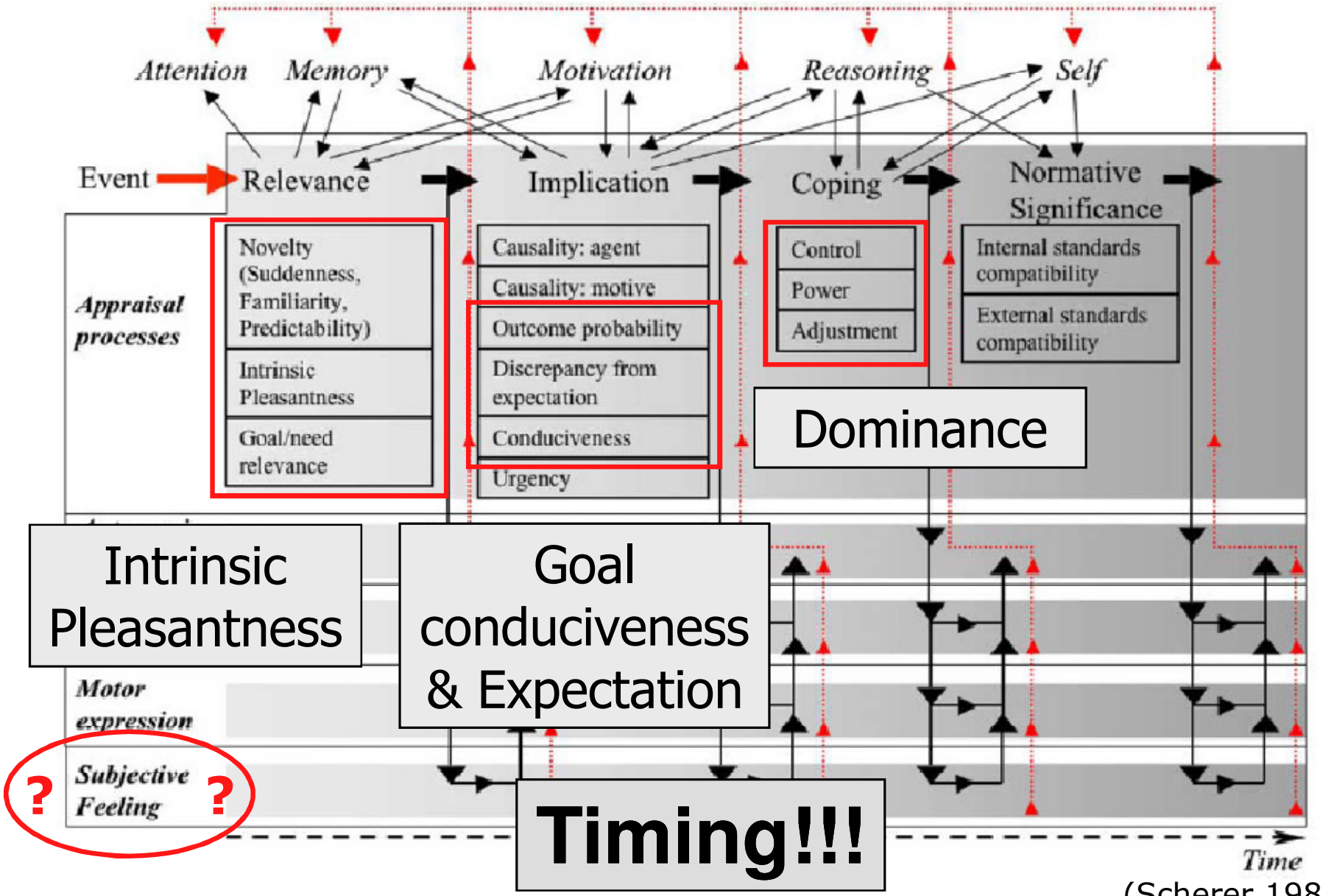
Attempts to implement the OCC-model only moderately successful

OCC-model best suited to *reason about* emotions

The non-cognitive emergence of emotions is neglected

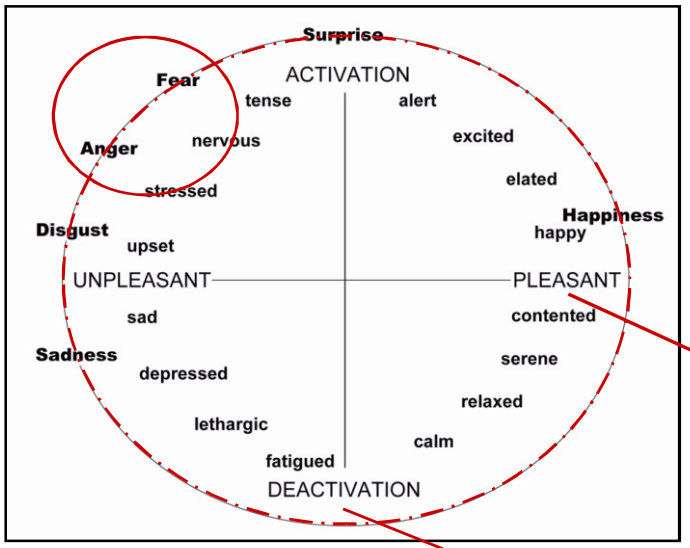
(E.g.: The experience of relaxation when sitting in front of a warm oven.)

Component Process Model



(Scherer 1984/2006)

Core Affect & PAD space



(Russel & Feldmann Barrett 1999)

Assumption underlying "Core Affect":

→ Emotions not identifiable by distinct categories from the start

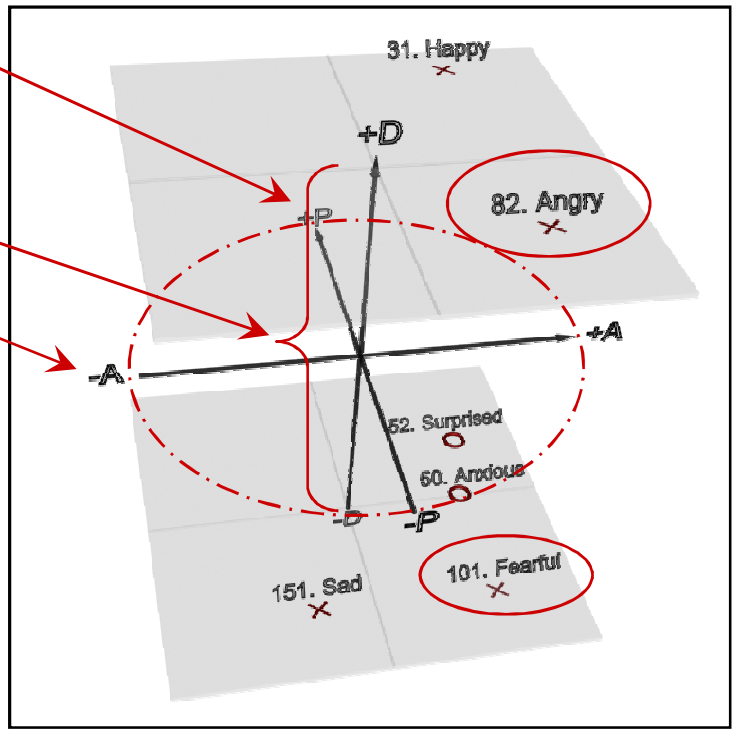
→ "Circumplex model of Core Affect" (Pleasentness & Activation)

Problem: "Fear" and "Anger" close together!

PAD space

Term	Pleasure		Arousal		**Dominance**	
	Mean	SD	Mean	SD	Mean	SD
31. Happy	.81	.21	.51	.26	.46	.38
50. Anxious	.01*	.45	.59	.31	-.15*	.32
52. Surprised	.40	.30	.67	.27	-.13*	.38
82. **Angry**	-.51	.20	.59	.33	**.25**	.39
101. **Fearful**	-.64	.20	.60	.32	**-.43**	.30
151. Sad	-.63	.23	-.27	.34	-.33	.22

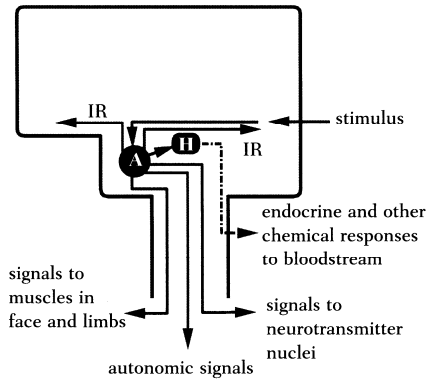
(Russel & Mehrabian 1977)



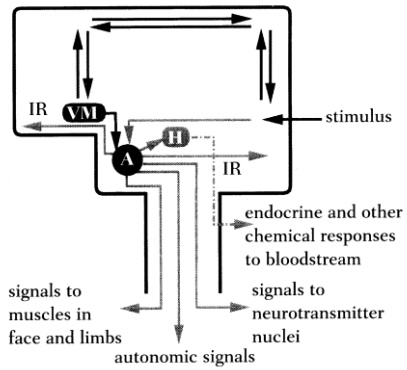
Neurobiological background

Primary and Secondary emotions:

1. Primary emotions (fear, anger, joy, ...):
   - fast, hard-wired stimulus response patterns
   - trigger fight-or-flight behaviors
   - ontogenetically earlier types of emotion



2. Secondary emotions (hope, relief, shame, ...):
   - lead to cognitively elaborated, deliberative behaviors
   - are based on memories and expectations
   - "social emotions" developed during infancy
   - "utilize the machinery of primary emotions"



(Damasio 1994; Holodynski & Friedlmeier 2005)

Summary and conclusions

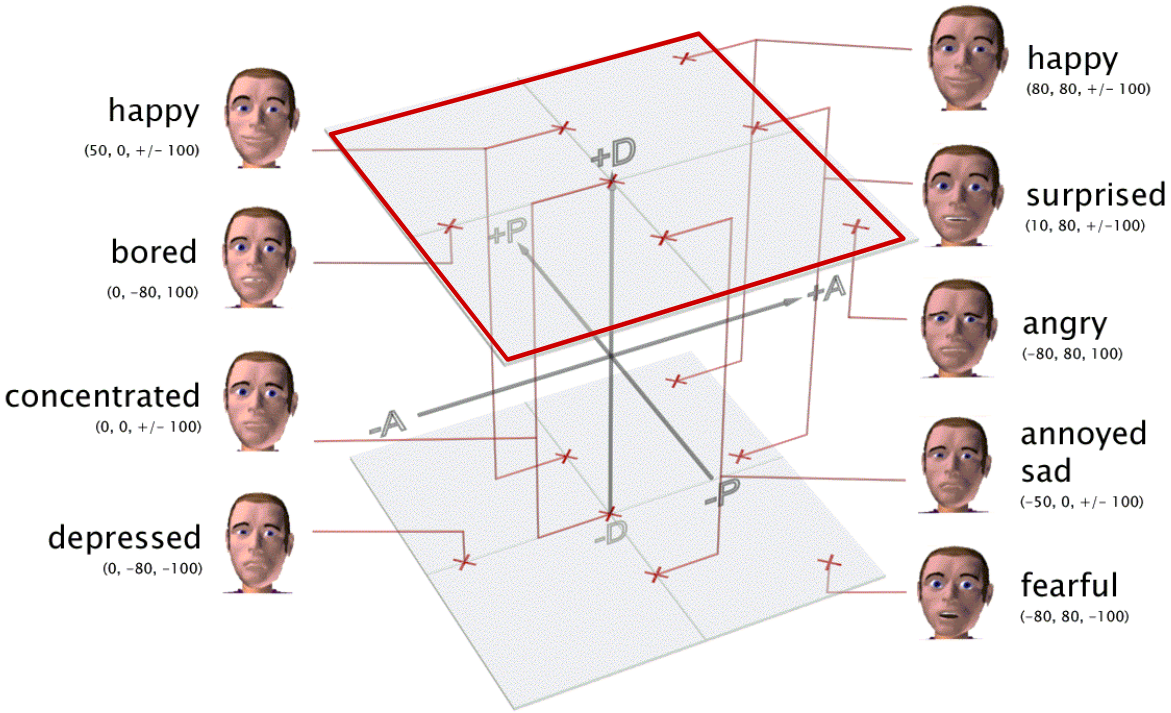
Primary and secondary emotions in context:

1. Primary emotions:
   - No memory, no expectations, no higher-order cognition
   - "Intrinsic pleasantness" (Scherer) and "Proto-affect" (Ortony et al.) used to realize "Gefühlsverlauf" (Wundt 1864)
   - Elicitation of primary emotions in PAD space (Russel & Mehrabian)
2. Secondary emotions:
   - product of conscious appraisal based on memory, expectations and goal-conduciveness (Scherer)
   - they "utilize the machinery of primary emotions"
     → also influence emotion dynamics
   - are product of ontogenetical development (Holodynski & Friedlmeier)
     → cannot be directly elicited in PAD space

Primary emotions: Max in the museum

Seven facial expressions for eight primary emotions (plus non-emotional state "concentrated")

in PAD space

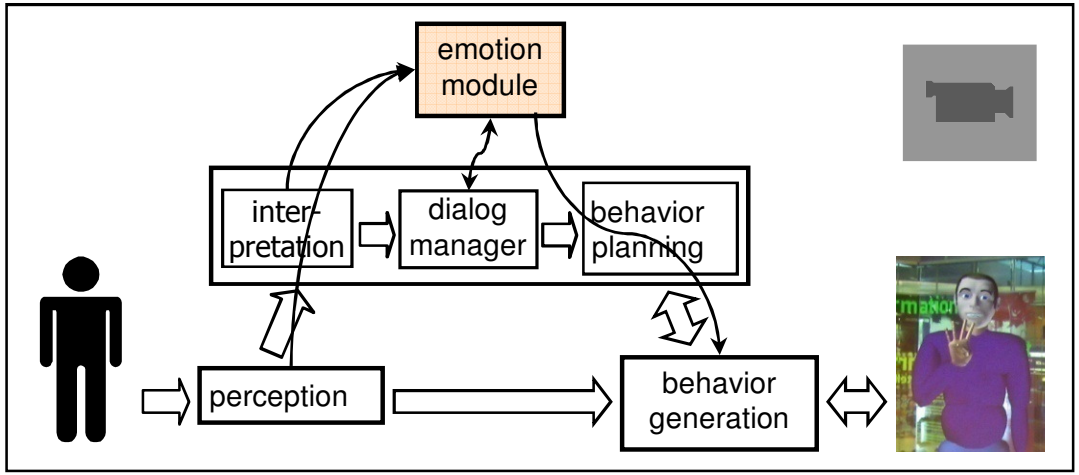


Direct elicitation of primary emotions

in the museum
(Gesellensetter 2004)

based on dynamics of emotional impulses

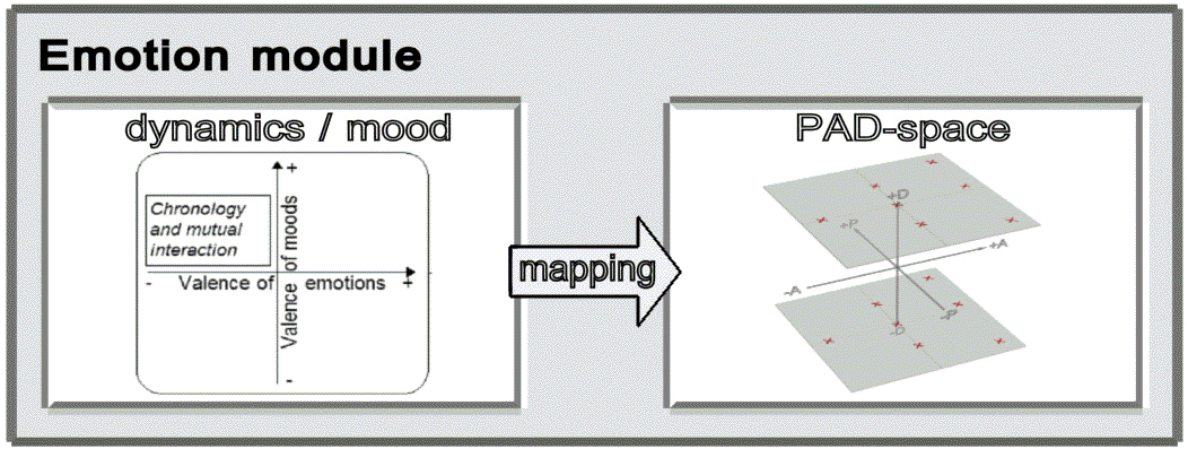
→ MAX is always dominant!



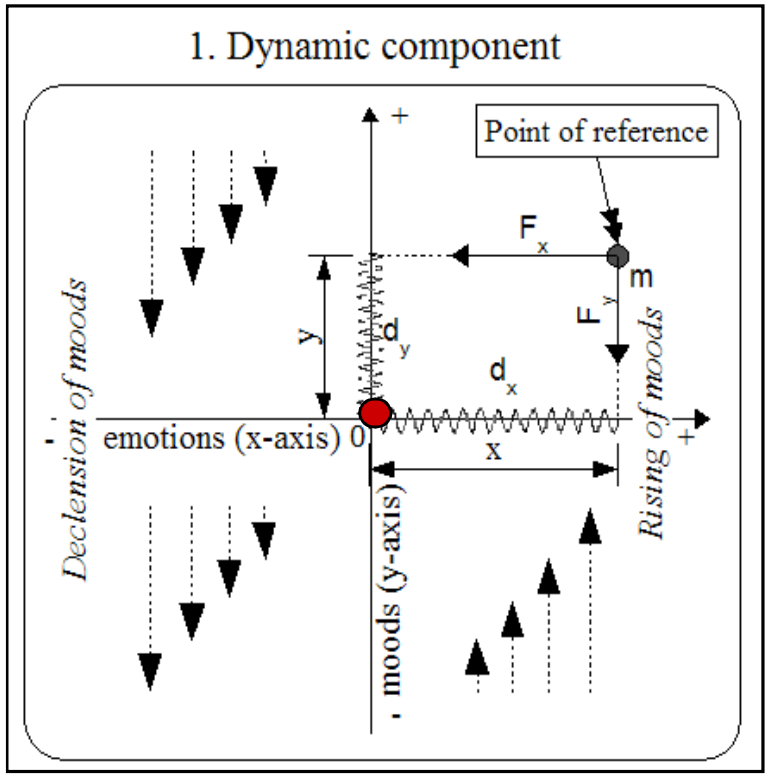
(Becker, Kopp, Wachsmuth 2004)

The emotion module (2004)

Conceptual linkage of emotion and mood



Elicitation of primary emotions based on a distance metric



- Valence of emotion (x) coupled with valence of mood (y) (see vertical arrows)
- Simulation of two spiral springs for emotion dynamics ($F_{(t)} = -d * x_{(t-1)}$)

Mapping into Pleasure-Arousal space:

$$K(x_t, y_t, z_t, t) = (p(x_t, y_t), a(x_t, y_t), d(t)),$$

$$with \quad p(x_t, y_t) = \frac{1}{2} \bullet (x_t + y_t) \quad and \quad a(x_t, z_t) = |x_t| + z_t$$

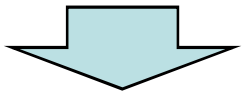
→Assures mood-congruent emotions

(Becker, Kopp, Wachsmuth 2004)

Open questions

What about..

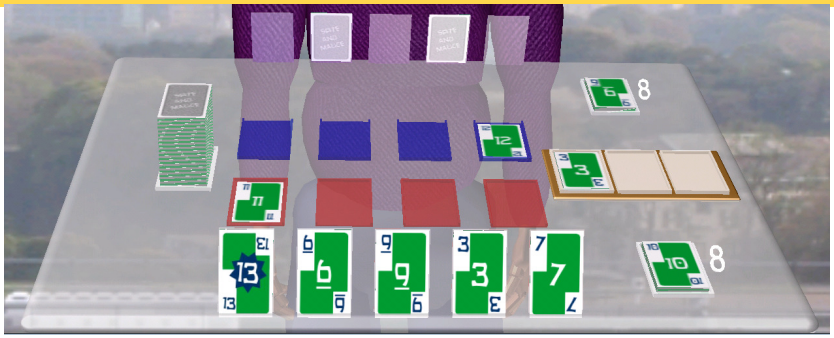
- Cognitively elaborated, secondary emotions (relief, hope, …)?
  →The WASABI architecture
- Determining the value of "Dominance"?
  →The Skip-Bo scenario
- Empirical evaluation of "believability"?



MAX playing Skip-Bo with primary emotion simulation

Primary emotions in the Skip-Bo scenario

Perceiving & expressing emotions



Thanks to JSPS and Prof. Prendinger

- Emotion recognition based on biometrical sensors (Skin Conductivity & Electromyography) provided by Prof. Prendinger
- Dominance adjusted according to game state

  → Max might fear to lose the game
- Empirical study conducted at NII, Japan, 2005
  - 32 participants
  - four conditions: Non-emotional, self-centered emotional, negative empathic, positive empathic

(Becker, Prendinger, Ishizuka & Wachsmuth 2005)

Empirical study: Results

For a competitive scenario!

1. Questionnaire:
   - An empathic MAX is judged more "human" than a non-empathic MAX
     → emotions increase believability of MAX
2. Biometrical data analysis:
   - Skin Conductance (SC): A negative empathic MAX is less stressful than a positive empathic MAX
     → negative emotions beneficial!
   - Electromyography: A negative empathic MAX induces negatively valenced emotions
     → emotional contagion, just as among humans!

Emotion dynamics simulation enhances believability of MAX

(Prendinger, Becker & Ishizuka 2006)

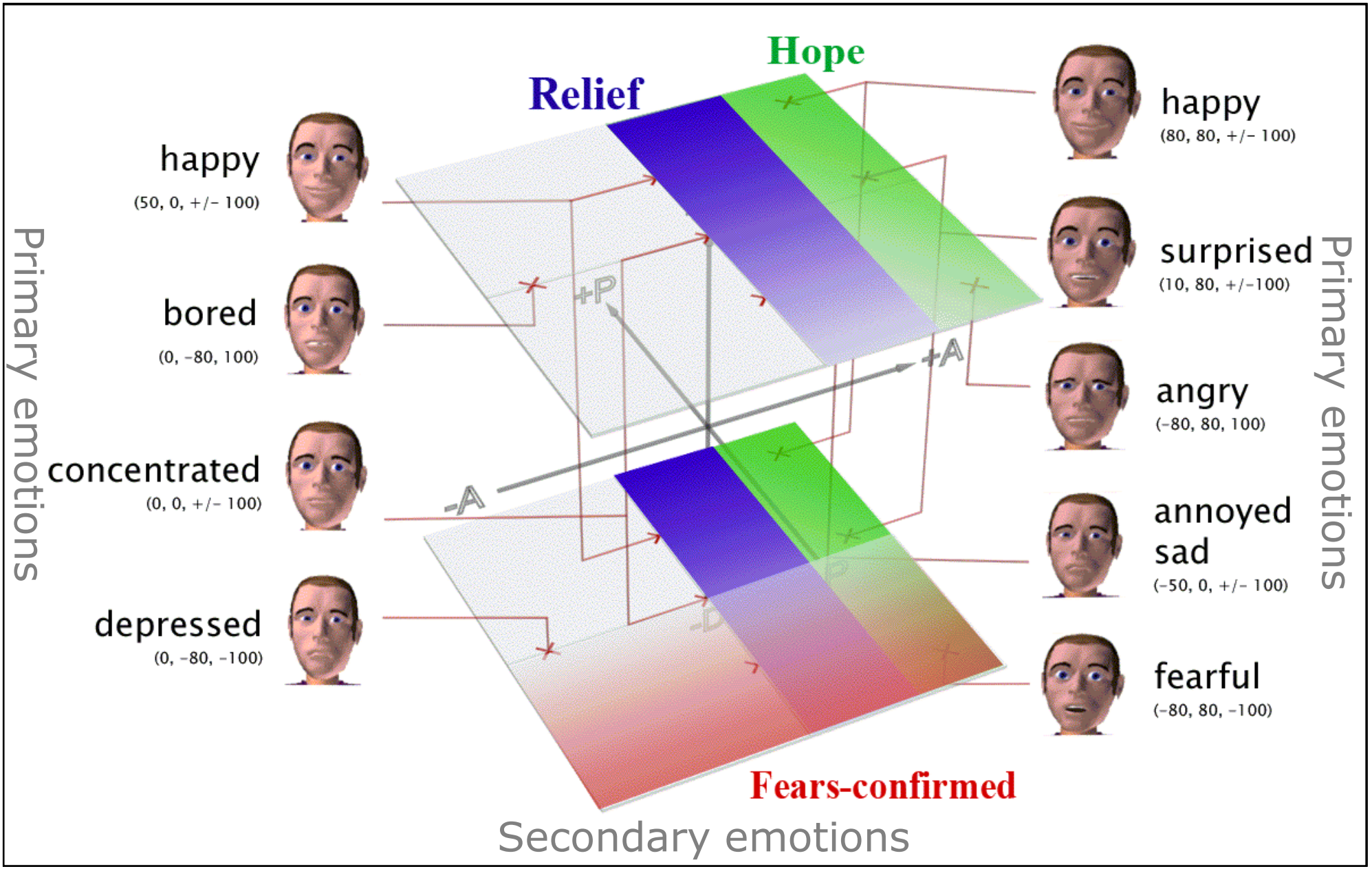
Integrating secondary emotions

Three prospect-based emotions as secondary emotions:

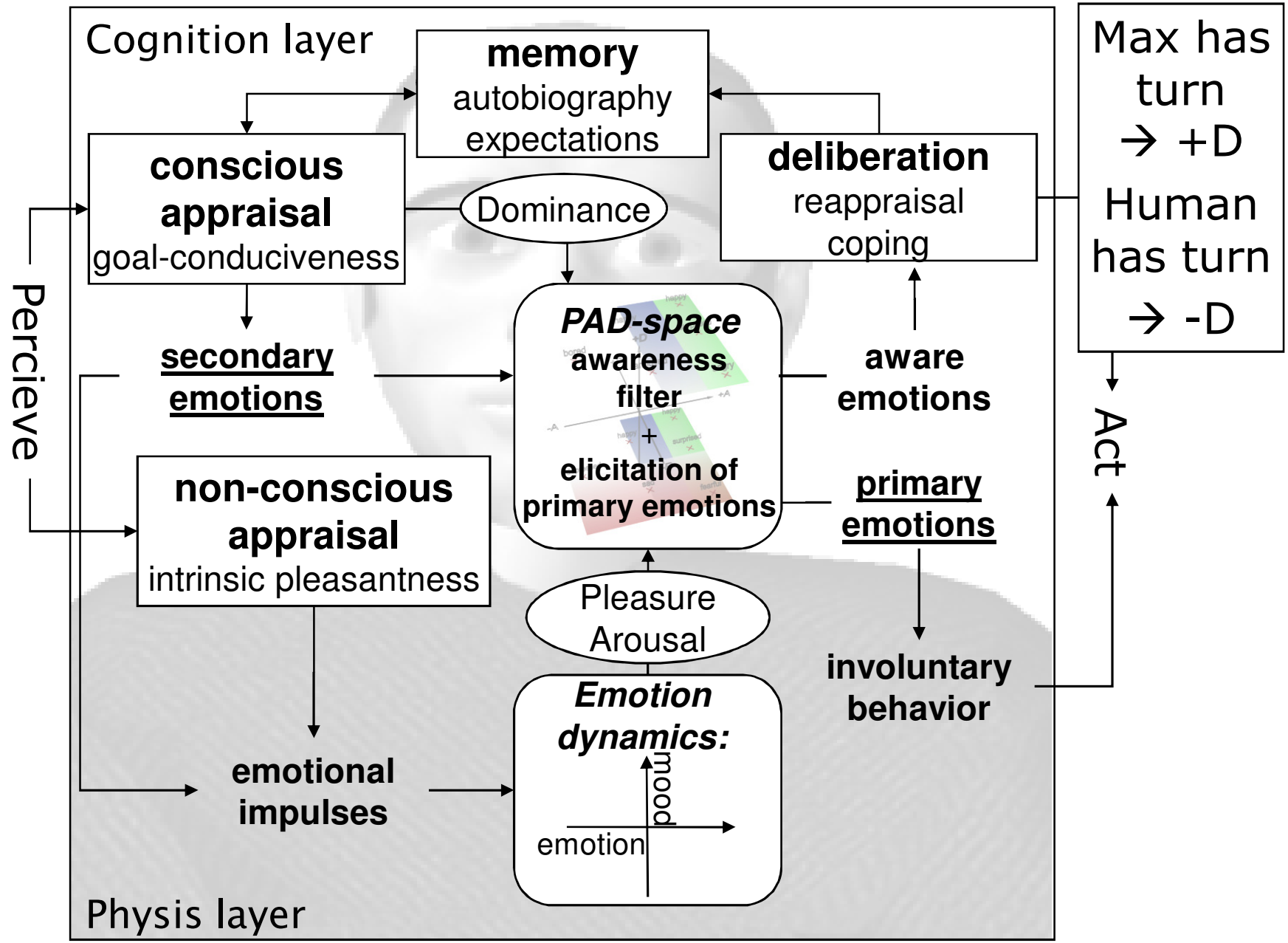
1. HOPE:
   MAX hopes that the human player plays one of her stock cards or her main card, because MAX could play his own main card afterwards.
2. FEARS-CONFIRMED:
   MAX sees his fears confirmed, because the human player played a card that MAX was afraid of before.
3. RELIEF:
   MAX realizes that the human player did NOT play a card that MAX was afraid could have been played.

Aspects of their connotative meaning represented in PAD space
→ ensuring mood-congruent elicitation
→ enables calculation of their awareness likelihood

Primary and secondary emotions in PAD space



The WASABI architecture



PAD space: Implementation example

Cognition: The prosp
undesirable event wa

happy
(50, 0, +/- 100)

bored
(0, -80, 100)

concentrated
(0, 0, +/- 100)

depressed
(0, -80, -100)

happy
(80, 80, +/- 100)

surprised
(10, 80, +/-100)

angry
(-80, 80, 100)

annoyed
sad
(-50, 0, +/- 100)

fearful
(-80, 80, -100)

"I was already afraid of that!"

Fears-confirmed

→ trigger **Fears-confirmed**

awareness likelihood = (0.3 * fearful, 0.2 * sad, 0.6 * **Fears-confirmed**)

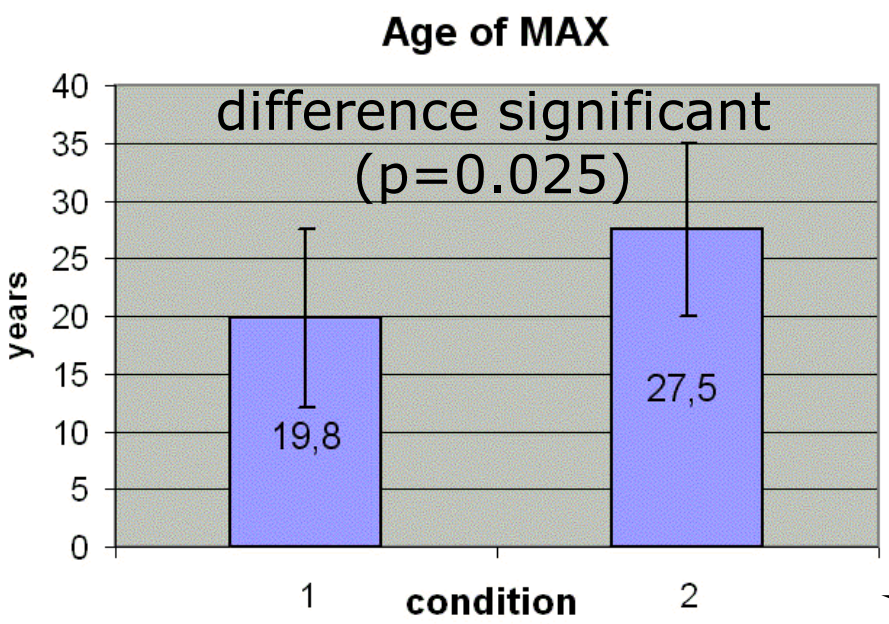
Empirical evidence

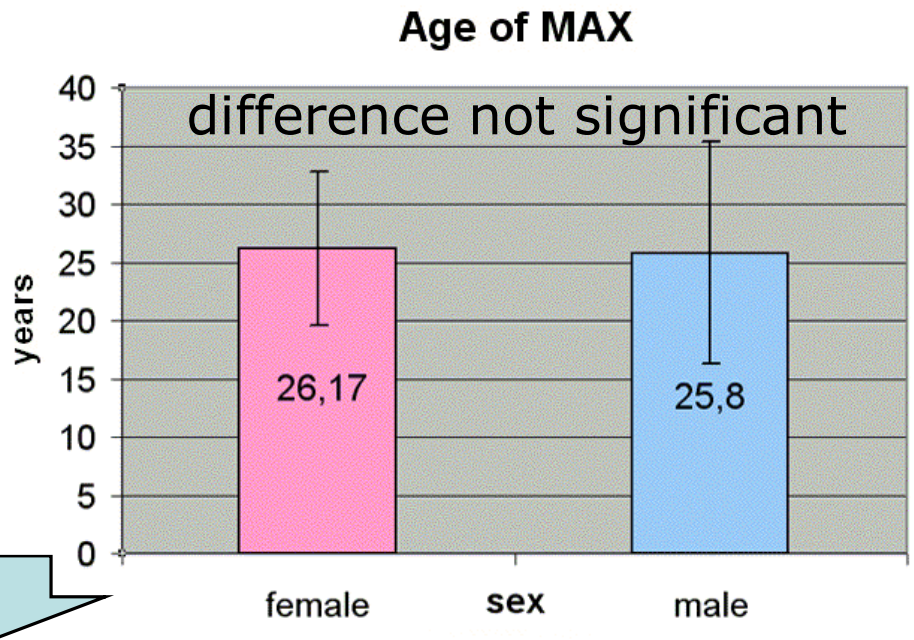
Two conditions:

1. Max only with primary emotions (N=11)
2. Max with primary and secondary emotions (N=12)

Question after playing (among others):

"If MAX were a real human, how old would you judge him?"



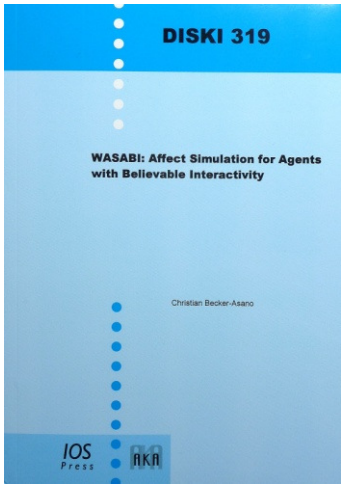


Secondary emotion simulation lets MAX appear older

(Becker-Asano & Wachsmuth 2008)

Summary

"I believe that the WASABI architecture is a helpful model to understand how the dynamic interplay of a human's body and mind together with his past experiences and future expectations sometimes turns 'cold' cognitions into 'hot' affective states."



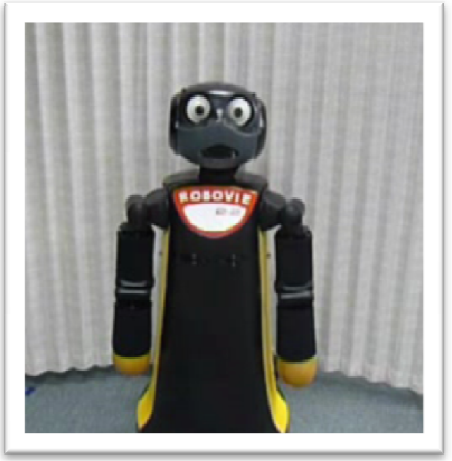
(Becker-Asano 2008)

Future perspectives & ongoing work

- From the virtual to the physical agent:
  - Expressiveness
  - Laughing
- Post-doctoral fellow of JSPS and Humboldt foundation (2008 – 2010)
- Host Prof. Ishiguro, ATR, Kyoto

Thank you for your attention!

Robovie II



Robovie R2

Geminoid HI-1 (ATR, Japan)