Noname manuscript No.
(will be inserted by the editor)

Affective computing with primary and secondary emotions in a virtual human

Christian Becker-Asano · Ipke Wachsmuth

Received: date / Accepted: date

Abstract We introduce the WASABI[1] Affect Simulation Architecture, in which a virtual human's cognitive reasoning capabilities are combined with simulated embodiment to achieve the simulation of primary and secondary emotions. In modeling primary emotions we follow the idea of "Core Affect" in combination with a continuous progression of bodily feeling in three-dimensional emotion space (PAD space), that is subsequently categorized into discrete emotions. In humans, primary emotions are understood as onto-genetically earlier emotions, which directly influence facial expressions. Secondary emotions, in contrast, afford the ability to reason about current events in the light of experiences and expectations. By technically representing aspects of each secondary emotion's connotative meaning in PAD space, we not only assure their mood-congruent elicitation, but also combine them with facial expressions, that are concurrently driven by primary emotions. Results of an empirical study suggest that human players in a card game scenario judge our virtual human MAX significantly older when secondary emotions are simulated in addition to primary ones.

1 Introduction

Researchers in the field of embodied agents (Cassell, Sullivan, Prevost & Churchill 2000; Prendinger & Ishizuka 2004) build anthropomorphic systems, which are employed in different interaction scenarios that afford communicative abilities of different style and complexity. As these agents comprise an increasing number of sensors as well as

This article is an extended version of Becker-Asano & Wachsmuth (2008).

Christian Becker-Asano
Faculty of Technology, University of Bielefeld, 33594 Bielefeld, Germany
E-mail: cbecker@techfak.uni-bielefeld.de
Present address:
ATR, 2-2-2 Hikaridai, Keihanna Science City, Kyoto 619-0288, Japan

Ipke Wachsmuth
Faculty of Technology, University of Bielefeld, 33594 Bielefeld, Germany
E-mail: ipke@techfak.uni-bielefeld.de

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

actuators together with an increase in expressive capabilities, they need to be able to process social cues in face-to-face communication in order to fulfill a human attendee's expectations.

One factor in social interaction is the ability to deal with the affective dimension appropriately. Therefore researchers in the growing field of "Affective Computing" (Picard 1997) discuss ways to derive human affective states from various intrusive and non-intrusive sensors. With regard to the expressive capabilities of embodied agents the integration of emotional factors influencing bodily expressions is argued for. These bodily expressions include, e.g., facial expression, body posture, and voice inflection and all of them must be modulated in concert to synthesize coherent emotional behavior.

In this work, scientific inquiry and engineering are closely interwoven, by taking a cognitive modeling approach (building on both symbolic and dynamic system paradigms). Creating an artificial system that reproduces certain aspects of a natural system can help us understand internal mechanisms that lead to particular effects. With the WASABI architecture we present our attempt to exploit different findings and conceptions of emotion psychology, neurobiology, and developmental psychology in a fully-implemented computational system. It is based on the simulation of an emotion dynamics in three-dimensional emotion space, which has proven beneficial to increase the lifelikeness and believability of our virtual human MAX in two different interaction scenarios (Becker, Kopp & Wachsmuth 2007). With this emotion dynamics simulation, however, we were limited to a class of rather simple emotions, that are similar to Damasio's conception of primary emotions. In the WASABI architecture we additionally use our agent's cognitive reasoning abilities to model the mood-congruent elicitation of secondary emotions as well.

In the following section the psychological as well as neurobiological background is described with respect to the distinction of primary and secondary emotions. Subsequently we give an overview of related work in the field of Affective Computing. The WASABI architecture is presented in Section 3. We explain how nine primary emotions together with three secondary emotions—namely the prospect-based emotions *hope*, *fears-confirmed*, and *relief*—were integrated in such a way that their mood-congruent elicitation can be guaranteed. We conclude by discussing results of a first empirical study on the effect of secondary emotion simulation in a card game scenario.

2 Background and related work

We will first introduce the psychological and neurobiological background before an overview of related work in affective computing is given.

2.1 Psychological and neurobiological background

Following the work of Scherer (1984) the psychological construct labeled emotion can be broken up into the component of cognitive appraisal, the physiological component of activation and Arousal, the component of motor expression, the motivational component, and the component of subjective feeling state. We follow the distinction of a cognitive appraisal component and a physiological component in our computational simulation of affect.[2] We also account for the motivational component, because our agent's reasoning capabilities are modeled according to the belief-desire-intention approach to modeling rational behavior (Rao & Georgeff 1991). We further believe that modeling the dynamic interaction between a cognitive and a physiological component is a promising first step toward the goal of computationally realizing subjective feelings.

Recently, psychologists started to investigate "unconscious processes in emotions" (Scherer 2005) and Ortony at al. discuss levels of processing in "effective functioning" by introducing a distinction between "emotions" and "feelings" (Ortony, Norman & Revelle 2005). They understand feelings as "readouts of the brain's registration of bodily conditions and changes" whereas "emotions are interpreted feelings" (Ortony et al. 2005, p. 174) and propose three different levels of information processing, which are compatible with Scherer's three modes of representation (Scherer 2005).

According to Ortony and colleagues, lower levels have to contribute in order to experience "hot" emotions such that "cold, rational anger" could be solely the product of the cognitive component "without the concomitant feeling components from lower levels." (Ortony et al. 2005, p. 197) A purely primitive feeling of fear, on the contrary, also lacks the necessary cognitive elaboration to become a hot emotion. These psychological considerations are compatible with LeDoux's distinction of a low and a high road of fear elicitation in the brain (LeDoux 1996) and Damasio's assumption of bodily responses causing an "emotional body state" (Damasio 1994, p. 138) that is subsequently analyzed in the thought process.

Damasio distinguishes at least two classes of emotions, namely, primary and secondary emotions (Damasio 1994), on the basis of his neurobiological findings. They are central to our simulation of affect and, thus, introduced next.

The class of "primary emotions" (Damasio 1994) are supposed to be innate. According to Damasio they developed during phylogeny to support fast and reactive response behavior in case of immediate danger, i.e. basic behavioral response tendencies like "flight-or-fight" behaviors. In humans, however, the perception of the changed bodily state is combined with the object that initiated it resulting in a "feeling of the emotion" with respect to that particular object (Damasio 1994, p. 132). Primary emotions are also understood as prototypical emotion types which can already be ascribed to one-year-old children (Damasio 2003). Furthermore, they are comparable to the concept of "Core Affect" (Feldman Barrett 2005; Russell & Feldmann Barrett 1999), which is based on the assumption that emotions cannot be identified by distinct categories from the start. "Core Affect", however, is represented in two-dimensional emotion space of Pleasure/Valence and Arousal, which is sometimes extended by a third dimension (Gehm & Scherer 1988; Russell & Mehrabian 1977) labeled "Control/Dominance/Power" of connotative meaning (Osgood, Suci & Tannenbaum 1957) (as empirically shown by Mehrabian (1996), each PAD component can be related to personality traits). Originally, Wundt claimed that any emotion can be characterized as a continuous progression in such three-dimensional emotion space (Wundt 1863). The three-dimensional abstract emotion space is commonly referred to as PAD space.

Secondary emotions like "relief" or "hope" are assumed to arise from higher cognitive processes, based on an ability to evaluate preferences over outcomes and expectations. Accordingly, secondary emotions are acquired during onto-genesis through learning processes in the social context. Damasio uses the adjective "secondary" to re-

[2] The motor expression component has been realized as well, but is not discussed here (cf. (Becker et al. 2007) for details).

fer to "adult" emotions, which "utilize the machinery of primary emotions" (Damasio 1994, p. 137) by influencing the acquisition of "dispositional representations", which are necessary for the elicitation of secondary emotions. These acquired dispositional representations, however, are believed to be different from the "innate dispositional representations" underlying primary emotions. Furthermore, secondary emotions influence bodily expressions through the same mechanisms as do primary emotions.

2.2 Related work

El-Nasr, Yen & Ioerger (2000) present FLAME as a formalization of the dynamics of 14 emotions based on fuzzy logic rules that includes a mood value, which is continuously calculated as the average of all emotion intensities to provide a solution to the problem of conflicting emotions being activated at the same time. The idea of expectations is realized in FLAME by means of learning algorithms based on rewards and punishments. Although it was not integrated into the simulation of a virtual human, the mutual influence of emotion and mood is quite similar to the conception of emotion dynamics in the WASABI architecture.

Marsella & Gratch (2006) focus with their EMA model on the dynamics of emotional appraisal. They also argue for a mood value as an addend in the calculation of otherwise equally activated emotional states following the idea of mood-congruent emotions. Their framework for modeling emotions is considered to be the first fully implemented, domain-independent architecture for emotional conversational agents.

Marinier & Laird (2004) aim to combine the work of Marsella & Gratch (2006) with the findings of Damasio (1994). In later publications (Marinier & Laird 2007), however, Damasio's work is less central and they follow the ideas of Scherer (2001). Their central idea of "appraisal frames" is based on the EMA model (Marsella & Gratch 2006) and eleven of Scherer's sixteen appraisal dimensions are modeled for integration in the Soar cognitive architecture. They distinguish an "Active Appraisal Frame", which is the result of a momentary appraisal of a given event, from a "Perceived Appraisal Frame", which results from the combination of the actual mood and emotion frames. Thereby, they take Damasio's distinction between emotion and feeling into account—similarly to the conception underlying the WASABI architecture.

With their development of "Greta" Pelachaud & Bilvi (2003) are mainly concerned with believability of conversational agents. To guarantee that Greta's facial expressions are always consistent with the situational context, de Rosis, Pelachaud, Poggi, Carofiglio & de Carolis (2003) model Greta's "mind" based on the BDI-approach. This architecture's emotion model builds upon a "Dynamic Belief Network", which integrates the time dimension in the representation of uncertainty of beliefs. Ochs, Devooght, Sadek & Pelachaud (2006) present another BDI-based approach to implement OCC-based appraisal for Greta taking into account the socio-cultural context and integrating a computational model of emotion blending for facial expressions. All of these approaches to simulate emotions by means of BDI-based reasoning, however, have difficulties to deal with temporal aspects of the dynamics of emotions, which is central to our work.

Although André, Klesen, Gebhard, Allen & Rist (1999) start by distinguishing primary and secondary emotions, this idea is not taken up in their later publications (e.g. Gebhard, Klesen & Rist 2004). Gebhard (2005), recently, uses PAD space to derive a mood value from emotions resulting from OCC-based appraisal. Three-dimensional emotion spaces similar to PAD space are also used to drive the sociable robot “Kismet” (Breazeal 2003) and the humanoid robot WE-4RII (Itoh, Miwa, Nukariya, Zecca, Takanobu, Roccella, Carrozza, Dario & Takanishi 2006).

3 The WASABI Architecture

The WASABI architecture conceptualized here builds upon previous work on the simulation of emotion dynamics for the virtual human MAX (Becker, Kopp & Wachsmuth 2004) that has proven to support the agent’s believability in two different interaction scenarios (Becker et al. 2007; Becker-Asano, Kopp, Pfeiffer-Leßmann & Wachsmuth 2008). It was, however, limited to the simulation of primary emotions.

Accordingly, the WASABI architecture (Becker-Asano 2008) combines bodily emotion dynamics with cognitive appraisal in order to simulate infant-like primary emotions as well as cognitively elaborated (more adult) secondary emotions. In the following a technically suitable specification of the different concepts *emotion* and *mood* is derived from the theoretical background presented above.

Emotions are understood as current states with a specific quality and intensity, which are the outcome of complex neurophysiological processes for communication. The processes include neural activity of the brain as well as physiological responses of the body. One gets aware of one’s emotions in two cases: (1) if their awareness likelihood w exceeds a certain threshold (cf. Section 3.3) or (2) if one concentrates on the underlying processes by means of introspection.

Emotions can be classified into primary and secondary ones, but every emotion has either positive or negative valence of a certain value and compared to mood an emotion lasts significantly less long. The differences between primary and secondary emotions are conceptualized as follows:

– Secondary emotions are based on more complex data structures than primary ones. Accordingly, only some general aspects of secondary emotions (such as their respective valence components) are represented in PAD space.
– The appraisal of secondary emotions depends much more on the situational context and an agent’s memory than that of primary emotions. Thus, secondary emotions are more dependent on the agent’s cognitive reasoning abilities.
– The releasers of secondary emotions might be learned based on the history of primary emotions in connection with memories of events, agents, and objects.
– The agent’s facial expressions of primary emotions (cf. Figure 1) may accompany secondary emotions such that they do not necessarily need to be expressed by their own set of facial expressions.
– Secondary emotions also modulate the agent’s simulated embodiment, such as its general level of Arousal.
– The agent expresses its awareness of secondary emotions verbally.

Mood is modeled as a background state with a much simpler affective quality than emotions. In contrast to the model of Gebhard (2005) mood is not derived from PAD space, but modeled as an agent’s overall feeling of well-being on a bipolar scale of positive versus negative valence already before a mapping into PAD space is achieved. Any non-neutral mood is slowly regulated back to a neutral state of mood—much slower than it is the case for emotional valence. Accordingly, a mood's duration is in general longer than that of any emotion.

The described interconnectivity of mood and emotions results in an "emotion dynamics" (described in more detail in (Becker et al. 2007)), in which mood influences the elicitation of emotions in such a way that mood-congruency of emotions is achieved. This idea is empirically supported by Neumann, Seibt & Strack (2001), who found that individuals in a positive mood are less likely to experience negative emotions and vice versa.

3.1 Nine primary emotions

According to the above discussion, primary emotions (PE) are inborn affective states, which are triggered by reflexes in case of potentially harmful stimuli. They result in fast, reactive behavioral responses and, thus, are quite similar to the concept of proto-affect proposed by (Ortony et al. 2005). According to developmental psychology, young children express their (primary) emotions directly, because they did not yet internalize this process as in the case of adults (Holodynski & Friedlmeier 2005).

In our previous realization of emotion dynamics (Becker et al. 2004) this direct expression of primary emotions is achieved by implementing five of Ekman's six "basic emotions" (Ekman, Friesen & Ancoli 1980). In addition, the emotions "bored", "annoyed", and "depressed" as well as the non-emotional state "concentrated" are simulated. Each of the primary emotions is located in PAD space according to Table 1, for which the coordinates are derived from some of the values given in (Russell & Mehrabian 1977, p. 286ff)[3]. Based on a distance metric in this emotion space they are dynamically elicited concurrently to the agent's cognition, which is also responsible for *triggering* these emotions in such a way that their respective base intensities i_{pe} are temporarily set to 1.0 (see Section 3.3 for further details).

Table 1 Primary emotions in PAD space: The five "basic emotions" of Ekman (1999) are assigned to the corresponding facial expressions modeled in (Becker et al. 2004) whenever such a mapping is possible (cf. Figure 1) and additionally an individual base intensity i_{pe} is set for each primary emotion (see also Section 3.3)

Primary emotion		Facial expr. (Ekman)	PAD values	base intensity i_{pe}
1.	angry	anger (*anger*)	(80, 80, 100)	0.75
2.	annoyed	sad (*sadness*)	(-50, 0, 100)	0.75
3.	bored	bored (*none*)	(0, -80, 100)	0.75
4.	concentrated	neutral (*none*)	(0, 0, ±100)	0.75
5.	depressed	sad (*sadness*)	(0, -80, -100)	0.75
6.	fearful	fear (*fear*)	(-80, 80, 100)	0.25
7.	happy	happy (*happiness*)	(80, 80, ±100) (50, 0, ±100)	0.75
8.	sad	sad (*sadness*)	(-50, 0, -100)	0.75
9.	surprised	surprised (*surprise*)	(10, 80, ±100)	0.0

[3] Although Russell & Mehrabian (1977) locate the emotion "bored" at a different point in PAD space, we are confident that we might label an emotional state of very low arousal and neutral valence with "bored" as well, because together with "depressed" it remains the only one modeled to consist of low arousal.

The seven facial expressions of MAX corresponding to the eight primary emotions and the neutral state "concentrated" are shown in Figure 1. In case of high Pleasure Ekman's set of basic emotions only contains one obviously positive emotion, namely happiness. Thus, in the WASABI architecture (Becker-Asano 2008) this primary emotion covers the whole area of positive Pleasure regardless of Arousal or Dominance as it is located in PAD space four times altogether.

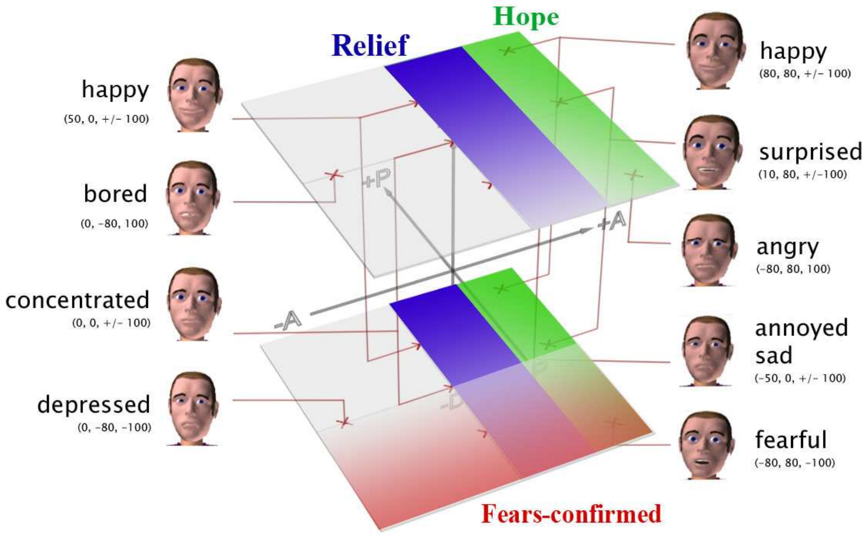


Fig. 1 The nine primary emotions of Table 1 as points together with the three secondary emotions of Table 2 as weighted areas in PAD space

3.2 Three secondary emotions

According to Damasio, the elicitation of secondary emotions involves a "thought process", in which the actual stimulus is evaluated against previously acquired experiences and online generated expectations.

The "prospect-based emotions" cluster of the OCC model of emotions (Ortony, Clore & Collins 1988) is considered here to belong to the class of secondary emotions, because their appraisal process includes the evaluation of events against experiences and expectations. This OCC cluster consists of six emotions (namely fear, hope, relief, disappointment, satisfaction, and fears-confirmed), of which *hope*, *fears-confirmed*, and *relief* are simulated in the WASABI architecture. The two secondary emotions *disappointment* and *satisfaction* could be implemented similarly, but have not yet been implemented in our demonstrator system due to time limitations.

The prospect-based emotion *fear* is obviously similar to the previously introduced primary emotion *fearful* (cf. Table 1). This similarity is accounted for by assigning a rather low base intensity of $i_6 = 0.25$ to the emotion *fearful*, such that MAX is less likely to get aware of this primary emotion, if it has not been *triggered* by the agent's cognition, making *fearful* a more cognition-dependent primary emotion (see also Section 4.1 for an example).

Table 2 The parameters of the secondary emotions *hope*, *fears-confirmed*, and *relief* for representation as weighted areas in PAD space (cf. Figure 1)

Area	(PAD values), intensity
HOPE	
high Dominance	(100, 0, 100), 0.6; (100, 100, 100), 1.0; (-100, 100, 100), 0.5; (-100, 0, 100), 0.1
low Dominance	(100, 0, -100), 0.6; (100, 100, -100), 1.0; (-100, 100, -100), 0.5; (-100, 0, -100), 0.1
FEARS-CONFIRMED	
low Dominance	(-100, 100, -100), 1.0; (0, 100, -100), 0.0; (0, -100, -100), 0.0; (-100, -100, -100), 1.0
RELIEF	
high Dominance	(100, 0, 100), 1.0; (100, 50, 100), 1.0; (-100, 50, 100), 0.2; (-100, 0, 100), 0.2
low Dominance	(100, 0, -100), 1.0; (100, 50, -100), 1.0; (-100, 50, -100), 0.2; (-100, 0, -100), 0.2

3.2.1 Hope

Ortony et al. (1988) describe *hope* as resulting from the appraisal of a prospective event. If the potential event is considered desirable for oneself, one is likely to be "pleased about the prospect of a desirable event" (Ortony et al. 1988, p. 110). The calculation of this emotion's awareness likelihood, however, is rather independent from these cognitive processes (see Section 3.3.2). This analysis provides the rationale for representing *hope* in PAD space in the following way:

- Pleasure: An agent is more likely to get aware of *hope* the more pleasurable he feels.
- Arousal: With respect to an agent's Arousal, *hope* is more likely elicited the higher the agent's Arousal value.
- Dominance: The awareness likelihood of *hope* is modeled to be independent of the agent's general level of Dominance.

To realize this distribution of awareness likelihood in case of hope two areas are introduced in Figure 1: one in the high Dominance plane and the other in the low Dominance plane. In Table 2 the exact values of the four corners of each of the two areas together with the respective base intensities in each corner are given for *hope*.

3.2.2 Fears-confirmed

According to Ortony et al., *fears-confirmed* is elicited when being "displeased about the confirmation of the prospect of an undesirable event." (Ortony et al. 1988, p. 110) With respect to its representation in PAD space the similarity to the primary emotion *fearful* is taken into account and the following decisions are taken:

- Pleasure: The awareness likelihood of *fears-confirmed* increases the less pleasurable the agent feels.

- Arousal: *fears-confirmed* is assumed to be independent of the agent's Arousal value.
- Dominance: *fears-confirmed* can only be perceived by the agent, when he feels submissive as in the case of *fearful*.

This distribution of awareness likelihood is realized in PAD space (cf. Figure 1) by introducing an area in the low Dominance plane (cf. Table 2 for the exact coordinates and intensities).

3.2.3 Relief

The secondary emotion *relief* is described as being experienced whenever one is "pleased about the disconfirmation of the prospect of an undesirable event." (Ortony et al. 1988, p. 110) Taking the similarity with Gehm and Scherer's "content" cluster into account Gehm & Scherer (1988), *relief* is represented in PAD space according to the following considerations:

- Pleasure: *relief* is more likely to become aware the more pleasurable the agent feels.
- Arousal: Only in case of relatively low Arousal levels the agent is assumed to get aware of the emotion *relief*.
- Dominance: The awareness likelihood of *relief* is considered to be independent from the agent's state of Dominance.

The awareness likelihood is represented in Figure 1 by two areas: one located in the high Dominance plane and the other in the low Dominance plane (cf. Table 2).

3.3 Emotion dynamics and awareness likelihood

The implementation of emotion dynamics is based on the assumption that an organisms natural, homeostatic state is characterized by emotional balance, which accompanies an agent's normal level of cognitive processing (Sloman, Chrisley & Scheutz 2005). Whenever an emotionally relevant internal or external stimulus is detected, however, its valence component serves as an emotional impulse, which disturbs the homeostasis causing certain levels of Pleasure and Arousal in the emotion module. Furthermore, a dynamic process is started by which these values are continuously driven back to the state of balance (cf. (Becker et al. 2007) for details).

The two valences are mathematically mapped into PAD space (cf. Figure 1) and combined with the actual level of Dominance, which is derived from the situational context in the cognition of the architecture. This process results in a course of a reference point in PAD space representing the continuously changing, bodily feeling state from which the awareness likelihoods of primary and secondary emotions are incessantly derived (see also Becker-Asano et al. 2008).

3.3.1 Awareness likelihood of primary emotions

The awareness likelihood of any of the nine primary emotions pe (cf. Table 1) depends on the distance between the actual PAD values and each primary emotion's PAD values (i.e. d_{pe} in Equation 1) with a smaller distance resulting in a higher awareness likelihood. When d_{pe} falls below Φ_{pe} units for a particular primary emotion pe, the calculation of its awareness likelihood w_{pe} is started according to Equation 1 until d_{pe}

falls below Δ_{pe} units in which case the likelihood w_{pe} equals the primary emotion's base intensity i_{pe}.

$$w_{pe} = (1 - \frac{d_{pe} - \Delta_{pe}}{\Phi_{pe} - \Delta_{pe}}) \cdot i_{pe}, \quad with \quad \Phi_{pe} > \Delta_{pe} \;\; \forall pe \in \{pe_1, \ldots, pe_9\} \tag{1}$$

In Equation 1, Φ_{pe} can be interpreted as the activation threshold and Δ_{pe} as the saturation threshold, which can be adjusted for every primary emotion $pe_n \in \{pe_1, \ldots, pe_9\}$ independently[4]. By setting a primary emotion's base intensity i_{pe} to 0.0 (as in the case of *surprised*, cf. Table 1) it needs to be *triggered* by the cognition before it might gain a non-zero awareness likelihood w_{pe}.

In case of primary emotions that are represented in PAD space more than once (i.e. concentrated, happy, and surprised; cf. Table 1) the representation with the minimum distance to the reference point is considered in Equation 1 for calculation of its awareness likelihood.

3.3.2 Awareness likelihood of secondary emotions

With representing the three secondary emotions in PAD space their mood-congruent elicitation can be assured, because the actual PAD values are also relevant for calculating every secondary emotion's awareness likelihood. In contrast to most primary emotions, all secondary emotions' *base intensities* are set to zero by default (as in case of *surprised*, cf. Table 1). Accordingly, every secondary emotion needs to be triggered by a cognitive process, before it gains the potential to get aware to the agent. Furthermore, a secondary emotion's *lifetime* parameter (set to 10.0 by default) together with its *decay function* (set to linear by default) are used to decrease its intensity over time until its *base intensity* of zero is reached again. As secondary emotions are represented in PAD space by four sided polygons (cf. Table 2) with separate intensities per vertex (i.e. corner), linear interpolation is used to calculate a secondary emotion's awareness likelihood (cf. (Becker-Asano 2008) for details).

3.4 Connecting cognition and embodiment

In Figure 2 the conceptual distinction of an agent's simulated embodiment and its cognition is presented and the different modules and components of the WASABI architecture are assigned to the corresponding layers.

To the left of Figure 2 the virtual human MAX perceives some (internal or external) stimulus. *Non-conscious appraisal* is realized by directly sending a small positive *emotional impulse* to the *Emotion dynamics*. This establishes the "low road" (LeDoux 1996) of primary emotion elicitation. For example, the presence of visitors in the museum (Becker et al. 2004) is interpreted as *intrinsically pleasant* following the ideas of Scherer (2001).

Another path resulting in *emotional impulses* begins with *conscious appraisal* of the perceived stimulus (cf. Figure 2, top left). This process resides in the *Cognition*, because it is based on the evaluation of goal-conduciveness of an event (Scherer 2001) and can be considered the "high road" of emotion elicitation (LeDoux 1996). Therefore, MAX exploits his BDI-based cognitive reasoning abilities to update his *memory* and

[4] The nine primary emotions are indexed according to Table 1 on page 6.

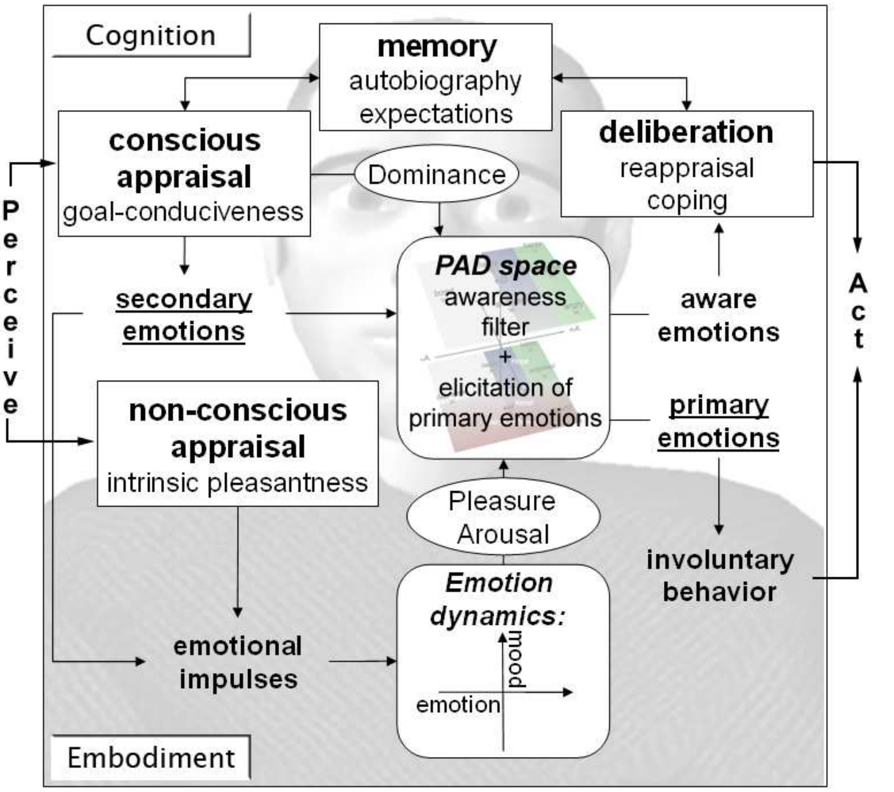


Fig. 2 The conceptual distinction of cognition and embodiment in the WASABI architecture. Any perceived stimulus is appraised by conscious and non-conscious processes in parallel leading to the elicitation of "emotional impulses". These drive the "emotion dynamics", which is part of the agent's virtual embodiment and from which mood, Pleasure, and Arousal are continuously derived. PAD space is used (1) to directly elicit primary emotions with a certain intensity and (2) to act as an "awareness filter", which ensures mood-congruency of both primary and secondary emotions. The resulting set of "aware emotions" are finally reappraised in the cognition before giving rise to deliberative actions.)

generate *expectations*. These deliberative processes not only enable MAX to derive his subjective level of *Dominance* from the situational and social context, but also propose cognitively plausible *secondary emotions*.

These *secondary emotions* are, however, first *filtered* in *PAD space*, before MAX might get *aware* of them (cf. Figure 2, middle). Independent of this "awareness filter", every cognitively plausible *secondary emotion* influences the *Emotion dynamics* component of the WASABI architecture, thereby modulating MAX's *Pleasure* and *Arousal* values, i.e. his simulated *Embodiment*. This influence is achieved by interpreting the valence component of any *secondary emotion* as an *emotional impulse* (cf. Figure 2, left). This way, *secondary emotions* "utilize the machinery of primary emotions" (Damasio 1994, p. 137), because they might result in the elicitation of mood-congruent *primary emotions*, which—in the WASABI architecture—drive MAX's facial expressions *involuntarily*. Furthermore, as the *Pleasure* and *Arousal* values are incessantly modulating MAX's *involuntary behaviors* (i.e. breathing and eye blinking) as well, even "unaware" *secondary emotions* have an effect on MAX's bodily state, which is expressed by these behaviors.

In combination with the actual level of *Dominance*, *primary emotions* are elicited by means of a distance metric in *PAD space*. As mentioned before, these primary emotions are directly driving MAX's facial expressions. Although this automatism might be considered unnatural for an adult, it has proven applicable and believable in the situational contexts in which MAX was integrated so far.

After the awareness filter has been applied, the resulting set of *aware emotions* consists of primary and secondary emotions, which have positive awareness likelihoods in that moment of the interaction. They are finally subject to further deliberation and reappraisal resulting in different coping behaviors (cf. Becker et al. 2007).

4 Application in the Skip-Bo scenario and first empirical evidence

The WASABI architecture was successfully used to extend the cognitive capabilities of MAX in the previously implemented Skip-Bo card game scenario (Becker, Prendinger, Ishizuka & Wachsmuth 2005; Prendinger, Becker & Ishizuka 2006). In this section its technical realization along an exemplary interaction are described and details of an empirical study are given, which yields promising results with regard to the effect of secondary emotion simulation.

4.1 Technical realization and example of an interaction

The virtual human MAX is based an a multi-agent system that enables us to encapsulate his cognitive abilities by devising specialized software agents (see Figure 3). They communicate with each other by passing messages.

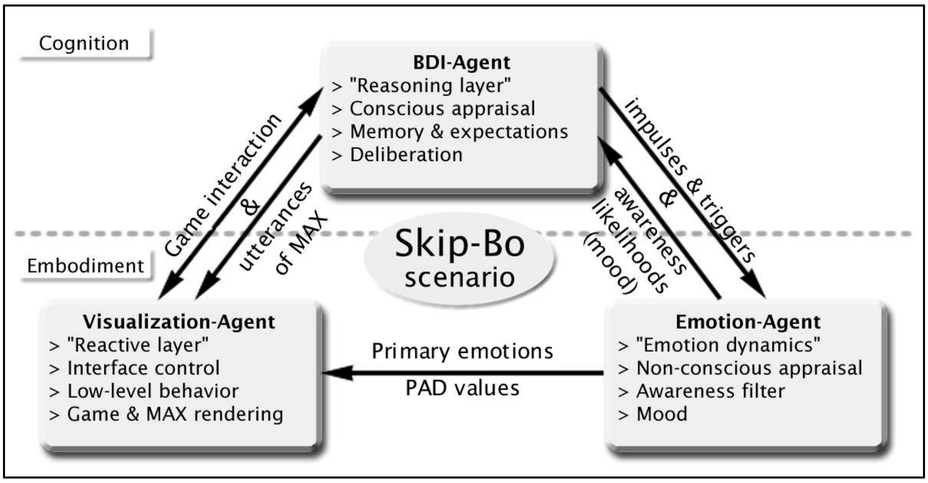


Fig. 3 The three most important software agents in the Skip-Bo scenario and their interconnection by means of message passing. Conceptually, the *BDI-Agent* realizes the cognition of MAX whereas his embodiment is simulated by two other software agents (cf. Figure 2)

The author's emotion dynamics simulation system is implemented as a so-called *Emotion-Agent* and acts in concert with a number of other agents (see Figure 3; cf. (Becker 2003; Becker-Asano 2008) for details). In the Skip-Bo scenario the *Emotion-Agent* receives emotional impulses from the *BDI-Agent*, which is continuously being updated with awareness likelihoods of primary and secondary emotions. The reasoning processes within the *BDI-Agent* also derive the actual state of Dominance from the context of the card game, such that MAX feels dominant whenever it is his turn and non-dominant, i.e. submissive, otherwise. Thus, whenever the human opponent fails to follow the rules of the game, MAX takes the turn to correct her and accordingly feels dominant until giving the turn back to the human. Concurrently, the *BDI-Agent* keeps the *Visualization-Agent* updated about the actual primary emotions and PAD values.

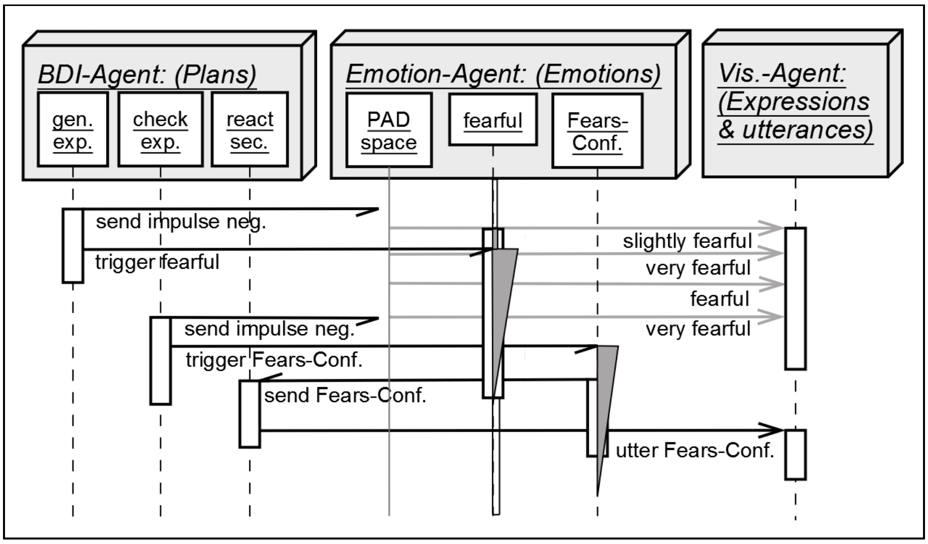


Fig. 4 Sequence diagram of an information flow between the software agents with the time-line from top to bottom

The sequence diagram in Figure 4 illustrates an example information flow within the WASABI architecture. The three agents *BDI-Agent*, *Emotion-Agent*, and *Visualization-Agent* ("Vis.-Agent") are represented as boxes in the top of Figure 4. In the top-left box, labeled *BDI-Agent*, the three plans *generate-expectation* ("gen. exp."), *check expectations* ("check exp."), and *react-to-secondary-emotion* ("react sec.") are rendered as three white rectangles to show their activity below[5]. The same rectangles are used to depict the *PAD space* as well as the emotions *fearful* and *Fears-Confirmed* ("Fears-Conf.") which all reside in the *Emotion-Agent*. The internal realization of the *Visualization-Agent* is not detailed here. In this example it only receives messages from the other agents, although in reality it also distributes information about the human player's interaction with the game interface by sending messages to the *BDI-Agent* (see Fig. 3).

[5] Pseudo-code representations as well as detailed descriptions of all BDI-plans can be found in (Becker-Asano 2008).

An exemplary sequence of message communication of these agents is shown in Figure 4 with the time-line from top to bottom. At first, the *generate-expectation* plan is called, e.g., after MAX played his last card. This plan, first, *sends* a *negative impulse* (“`send impulse neg.`”) to the *Emotion-Agent* thereby indirectly changing MAX’s emotional state in *PAD space* (cf. Section 3.3). Subsequently, while following the same plan, the primary emotion *fearful* is being *triggered* (“`trigger fearful`”) by the *BDI-Agent*, because MAX expects the human player to play an important card.

In the *Emotion-Agent*, however, the negative emotional impulse pushed the reference point in PAD space already close enough to the (not yet triggered) emotion *fearful* to let MAX experience *fear* with low intensity, because *fearful* has a slightly positive base intensity of 0.25 (cf. Table 1). In Figure 4 this non-zero standard intensity of *fearful* is indicated by a small double line along the dashed, vertical lifeline of *fearful*. Accordingly, “slightly fearful” is sent to the *Visualization-Agent* even before the *BDI-Agent* triggers the emotion *fearful*. Because the intensity of *fearful* in the *Emotion-Agent* abruptly changes with the incoming *trigger fearful* message, MAX’s emotional state changes from *slightly fearful* to *very fearful*. This sudden change in intensity is reproduced in Figure 4 by the two gray triangles drawn along each emotion’s lifelines and leads to a clear expression of fear in MAX’s face as shown in Figure 5(a).

The intensity of *fearful* decreases within the next ten seconds and the reference point changes its location in PAD space due to the implemented emotion dynamics. Thus, *very fearful* automatically changes to *fearful* (see right side of Figure 4) without any further *impulse* or *trigger* messages.

Next, the *check expectations* plan is activated in the *BDI-Agent* to check, if a human player’s action meets the previously generated expectations. In this example, the *BDI-Agent*, first, sends a *negative impulse* to the *Emotion-Agent*, thereby changing the reference point’s location in PAD space such that MAX gets *very fearful* again. This sequence of different emotion intensities (*slightly fearful*, *very fearful*, *fearful*, *very fearful*) is possible for every primary or secondary emotion, although it is exemplified only for *fearful* in Figure 4. It results from the dynamic interplay of lower-level emotional impulses and higher-level (cognitively triggered) changes of an emotion’s base intensity.

The *check expectations* plan then *triggers* the secondary emotion *Fears-Confirmed* (“`trigger Fears-Conf.`”) in the *Emotion-Agent*, thereby maximizing its base intensity. Together with the negatively valenced mood, *fears-confirmed* acquires a non-zero awareness likelihood, which is *sent* back to the *BDI-Agent* (“`send Fears-Conf.`”). The plan *react-to-secondary-emotion* is then executed to process the incoming message and results in an “`utter Fears-Conf.`” message, which is *sent* to the *Visualization-Agent* letting MAX produce an appropriate utterance (cf. Figure 5(b)).

A human opponent would possibly give the following description of this sequence[6]:

> After MAX ended his turn with playing a hand card to one of his stock piles, he seemed to realize within one or two seconds that I could now directly play one of my four stock pile cards, because his facial expression changed to fearful and he seemed to inhale sharply, which produced the characteristic sound of someone being afraid. When I then actually played that stock card, MAX admitted that he had been afraid of that before.

[6] A video of this and similar interactions is available for download at http://www.becker-asano.de/WASABI_MaxInCave_Page.avi.

(a) MAX *fears* that the human might play his stock pile card next and displays a fearful facial expression.

(b) MAX realizes that his *fears* just got *confirmed* and utters “Das hatte ich schon befürchtet!” (I was afraid of that!)

Fig. 5 MAX first *fearing* the human player to play a card and then seeing his *fears confirmed*

Similar dynamic interplays between conscious appraisals and bodily feelings are realized within the Skip-Bo scenario for the secondary emotions *relief* and *hope* as well (cf. Becker-Asano 2008).

4.2 Empirical evidence on the effect of secondary emotion simulation

Although our approach has not been subject to extended evaluation, some empirical evidence on the effect of secondary emotion simulation could already be gathered[7]. We derived the following hypothesis from the psychological findings presented in Section 2.1.

> **Hypothesis** MAX expressing primary and secondary emotions is judged older than MAX expressing only primary emotions.

As discussed in Section 2.1, secondary emotions are understood to be the product of ontogenetical development and children are less able to suppress their emotional expressions than adults. Accordingly, humans playing Skip-Bo against MAX with secondary emotions (i.e., driven by the WASABI architecture) should be expected to judge him older than those humans that play against a version of MAX which only simulates and directly expresses primary emotions (i.e., driven by the original emotion dynamics system of Becker (2003); cf. (Becker et al. 2004)).

4.2.1 Design

To test the above hypothesis the following two conditions[8] were designed:

[7] We can only present a summary of the empirical study here, but a detailed description can be found in (Becker-Asano 2008).

[8] Notably, the number of verbal utterances performed by MAX is likely to be higher in condition (2) than in condition (1). This difference, however, adds to the impression of MAX

(1) *Only primary emotions* condition: The emotion simulation is constrained to primary emotions and MAX expresses them directly by means of facial expressions and "affective sounds". He appraises the actions of the human player negatively and his own progress in the game positively. During his turn and while correcting a human's mistake he feels *dominant*, otherwise *submissive* (i.e. non-dominant).
(2) *Primary and secondary emotions* condition: Secondary emotions are simulated in addition to the setup of condition (1) and MAX expresses them verbally in case of positive awareness likelihood (cf. Section 4.1).

In order to model condition (1) the WASABI architecture is initialized such that:

- The three secondary emotions *hope*, *fears-confirmed*, and *relief* are not included.
- Every primary emotion has the same *saturation threshold* Δ_{pe} (0.2), *activation threshold* Φ_{pe} (0.64), and *base intensity* (1.0).

In effect, the simpler *emotion simulation system* of Becker (2003) is reproduced within the WASABI architecture for condition (1).

To realize condition (2) the emotion module is initialized according to Tables 1 and 2 (cf. Section 3.1, page 6ff.). Furthermore, for *surprised* the *saturation threshold* Δ_9 is set to 0.3 in order to increase the probability of MAX getting aware of his surprise, after this primary emotion was *triggered* by the cognition (cf. (Becker-Asano 2008) for details)[9].

4.2.2 Participants

Fourteen male and nine female university students voluntarily took part in the study and all but one of them were German. Their age ranged from 13 to 36 years and the average age was 23 years. Participants were randomly assigned to the conditions resulting in a total of 11 participants (six male and five female) for condition (1) and 12 participants (eight male and four female) for condition (2).

4.2.3 Procedure

Participants received written instructions of the card game (in German) with a screenshot of the starting condition and got the chance to ask clarifying questions about the gameplay before they entered a room with a three-sided large-screen projection system. Participants entered the room individually and were equipped with goggles for 3D viewing and a marker for the right hand. They were briefed about the experiment, in particular that they would play a competitive game. Then, the participants played a short introductory game against a non-emotional MAX, which allowed them to get used to the interface, and also gave them the chance to ask clarifying questions about the game. Each participant won this first game easily.

From now on, the experimenter remained visually separated from the participant only to supervise the experiment. After the game was reset manually, MAX welcomed

as a less child-like interaction partner in condition (2), because young children are also less good at expressing their feelings verbally.

[9] *Surprised* is the only primary emotion with a base intensity of 0.0 (cf. Table 1) such that it can only get aware to the agent after being triggered by the cognition. This is based on the assumption that surprised is only expressed in reaction to unexpected events and expectedness of events is checked within the *BDI-Agent*, i.e. the agent's cognition.

the participant verbally and asked him or her to play the first card. After the game was completed, participants were asked to fill in a questionnaire in German presented on the screen of another computer in a room next door.

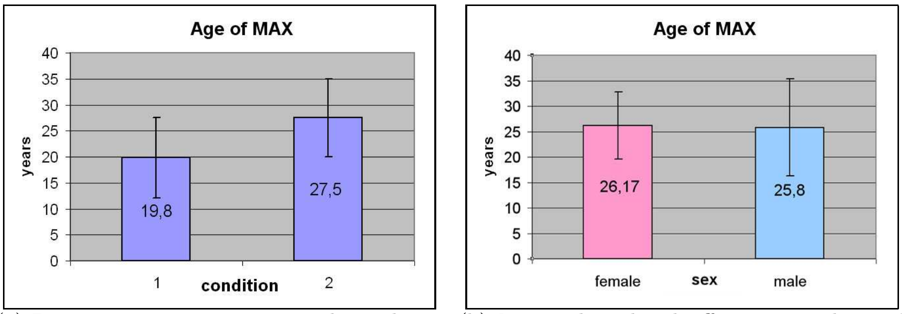
This questionnaire contained a total of 26 questions, of which 25 had already been asked in Japanese to the participants of a similar study in Japan (cf. (Prendinger et al. 2006)). To falsify our hypothesis we additionally asked the following question:

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

The participants were asked to fill in MAX's suspected age in years in a blank behind this question.

4.2.4 Results

The two mean values and standard deviations of the participants' answers to the above question are presented in Figure 6(a). In the *primary emotions only* condition (1) MAX was judged to be significantly younger (mean value 19.8 years, standard deviation 7.7) than in condition (2), in which secondary emotions were simulated as well (mean value 27.5, standard deviation 7.5). A two-tailed t-test assuming unequal variances results in p = 0.025.



(a) Between *primary emotions only* condition (1) and *primary and secondary emotions* condition (2) a significant difference occurred

(b) No gender-related effects were observed with regard to question 17b

Fig. 6 Mean values and standard deviations of the answers to question number 17b "If MAX were a real human, how old would you judge him to be?"

Because male and female participants were not evenly distributed among the two conditions (cf. Section 4.2.2), the answers of all nine female participants were compared to the answers of all 14 male participants regardless of the experimental condition. The mean values of these two groups did not differ significantly (cf. Figure 6(b)) letting us assume that no gender effects occurred. This result strengthens the supposition that the between-conditions difference can be interpreted to confirm the initial hypothesis.

5 Discussion and conclusion

We presented the WASABI architecture for mood-congruent simulation of primary and secondary emotions as it is integrated in, and makes use of, the overall cognitive architecture of the virtual human MAX. The simulation and direct expression of primary emotions is based on the idea to capture an agent's bodily feeling as a continuous progression in three-dimensional emotion space (i.e. PAD space), which is subsequently translated into weighted, primary emotions. Secondary emotions, in contrast, are understood as a class of emotions that require higher cognitive reasoning abilities and a certain sense of time, in that an agent has to be able to take experiences and expectations into account to generate prospect-based, secondary emotions. To also assure mood-congruency of secondary emotions, we capture aspects of their connotative meanings in PAD space as well by introducing weighted areas. Furthermore, to account for the decisive influence of cognitive processes in the elicitation of secondary emotions, they can gain a certain awareness likelihood in PAD space of the agent's virtual embodiment, only after having been triggered by cognitive processes.

The simulation of secondary emotions was exemplified by integrating three prospect-based emotions into the WASABI architecture. We believe, however, that other high-level, secondary emotions could be simulated similarly.

Finally, we reported on an empirical study conducted to answer a question derived from developmental psychology, namely, if the additional simulation of secondary emotions lets our virtual human MAX appear older within a playful interaction. Although the results of this study can be interpreted affirmatively, we acknowledge that the difference in MAX's use of spoken language between conditions may have compromised the results. Independent thereof, the practical applicability of the WASABI architecture to a well-defined, playful interaction scenario was demonstrated successfully.

In conclusion, we believe that the WASABI architecture is a helpful model to understand how the dynamic interplay of body and mind together with past experiences and future expectations sometimes turns "cold" cognitions into "hot" affective states.

6 Acknowledgments

We would like to express our cordial thanks to the two anonymous reviewers for providing us with very detailed and helpful comments.

References

André, E., Klesen, M., Gebhard, P., Allen, S., & Rist, T. (1999). Integrating models of personality and emotions into lifelike characters. In *Proceedings International Workshop on Affect in Interactions - Towards a New Generation of Interfaces*, (pp. 136–149).

Becker, C. (2003). Simulation der Emotionsdynamik eines künstlichen humanoiden Agenten. Master's thesis, University of Bielefeld.

Becker, C., Kopp, S., & Wachsmuth, I. (2004). Simulating the emotion dynamics of a multimodal conversational agent. In *Affective Dialogue Systems*, LNAI 3068, (pp. 154–165). Springer.

Becker, C., Kopp, S., & Wachsmuth, I. (2007). Why emotions should be integrated into conversational agents. In T. Nishida (Ed.), *Conversational Informatics: An Engineering Approach* chapter 3, (pp. 49–68). Wiley.

Becker, C., Prendinger, H., Ishizuka, M., & Wachsmuth, I. (2005). Evaluating Affective Feedback of the 3D Agent Max in a Competitive Cards Game. In *Affective Computing and Intelligent Interaction*, LNCS 3784, (pp. 466–473). Springer.

Becker-Asano (2008). *WASABI: Affect Simulation for Agents with Believable Interactivity*. PhD thesis, Faculty of Technology, University of Bielefeld. IOS Press (DISKI 319).

Becker-Asano, C., Kopp, S., Pfeiffer-Leßmann, N., & Wachsmuth, I. (2008). Virtual Humans Growing up: From Primary Toward Secondary Emotions. *KI Zeitschift (German Journal of Artificial Intelligence)*, *1*, 23–27.

Becker-Asano, C. & Wachsmuth, I. (2008). Affect simulation with primary and secondary emotions. In Prendinger, H., Lester, J., & Ishizuka, M. (Eds.), *Intelligent Virtual Agents (IVA 08)*, LNAI 5208, (pp. 15–28). Springer.

Breazeal, C. (2003). Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, *59*, 119–155.

Cassell, J., Sullivan, J., Prevost, S., & Churchill, E. (2000). *Embodied Conversational Agents*. Cambridge, MA: The MIT Press.

Damasio, A. (1994). *Descartes' Error, Emotion Reason and the Human Brain*. Grosset/Putnam.

Damasio, A. (2003). *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*. Harcourt.

de Rosis, F., Pelachaud, C., Poggi, I., Carofiglio, V., & de Carolis, B. (2003). From Greta's mind to her face: modelling the dynamics of affective states in a conversational embodied agent. *International Journal of Human-Computer Studies. Special Issue on "Applications of Affective Computing in HCI"*, *59*, 81–118.

Ekman, P. (1999). Facial expressions. In *Handbook of Cognition and Emotion* chapter 16, (pp. 301–320). John Wiley & Sons.

Ekman, P., Friesen, W., & Ancoli, S. (1980). Facial sings of emotional experience. *Journal of Personality and Social Psychology*, *29*, 1125–1134.

El-Nasr, M. S., Yen, J., & Ioerger, T. R. (2000). FLAME - Fuzzy Logic Adaptive Model of Emotions. *Autonomous Agents and Multi-Agent Systems*, *3*(3), 219–257.

Feldman Barrett, L. (2005). Feeling is perceiving: Core affect and conceptualization in the experience of emotion. In *The unconscious in emotion* chapter 11, (pp. 255–284). Guilford Press.

Gebhard, P. (2005). ALMA - A Layered Model of Affect. In *Autonomous Agents & Multi Agent Systems*, (pp. 29–36).

Gebhard, P., Klesen, M., & Rist, T. (2004). Coloring multi-character conversations through the expression of emotions. In *Affective Dialogue Systems*, (pp. 128–141).

Gehm, T. L. & Scherer, K. R. (1988). Factors determining the dimensions of subjective emotional space. In K. R. Scherer (Ed.), *Facets of Emotion* chapter 5. Lawrence Erlbaum Associates.

Holodynski, M. & Friedlmeier, W. (2005). *Development of Emotions and Emotion Regulation*. Springer.

Itoh, K., Miwa, H., Nukariya, Y., Zecca, M., Takanobu, H., Roccella, S., Carrozza, M., Dario, P., & Takanishi, A. (2006). Behavior generation of humanoid robots depending on mood. In *9th International Conference on Intelligent Autonomous Systems (IAS-9)*, (pp. 965–972).

LeDoux, J. (1996). *The Emotional Brain*. Touchstone. Simon & Schuster.

Marinier, R. & Laird, J. (2004). Toward a comprehensive computational model of emotions and feelings. In *International Conference on Cognitive Modeling*.

Marinier, R. P. & Laird, J. E. (2007). Computational modeling of mood and feeling from emotion. In *CogSci*, (pp. 461–466).

Marsella, S. & Gratch, J. (2006). EMA: A computational model of appraisal dynamics. In *Agent Construction and Emotions*.

Mehrabian, A. (1996). Analysis of the big-five personality factors in terms of the pad temperament model. *Australian Journal of Psychology*, *48*, 86–92.

Neumann, R., Seibt, B., & Strack, F. (2001). The influence of mood on the intensity of emotional responses: Disentangling feeling and knowing. *Cognition & Emotion*, *15*, 725–747.

Ochs, M., Devooght, K., Sadek, D., & Pelachaud, C. (2006). A computational model of capability-based emotion elicitation for rational agent. In *Workshop Emotion and Computing - German Conference on Artificial Intelligence (KI)*.

Ortony, A., Clore, G. L., & Collins, A. (1988). *The Cognitive Structure of Emotions*. Cambridge: Cambridge University Press.

Ortony, A., Norman, D., & Revelle, W. (2005). Affect and proto-affect in effective functioning. In J. Fellous & M. Arbib (Eds.), *Who needs emotions: The brain meets the machine* (pp. 173–202). Oxford University Press.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The Measurement of Meaning*. University of Illinois Press.

Pelachaud, C. & Bilvi, M. (2003). Computational model of believable conversational agents. In M.-P. Huget (Ed.), *Communication in MAS: background, current trends and future*. Springer-Verlag. To appear.

Picard, R. W. (1997). *Affective Computing*. Cambridge, MA: The MIT Press.

Prendinger, H., Becker, C., & Ishizuka, M. (2006). A study in users' physiological response to an empathic interface agent. *International Journal of Humanoid Robotics*, *3*(3), 371–391.

Prendinger, H. & Ishizuka, M. (2004). *Life-Like Characters. Tools, Affective Functions, and Applications*. Cognitive Technologies. Berlin Heidelberg: Springer Verlag.

Rao, A. & Georgeff, M. (1991). Modeling Rational Agents within a BDI-architecture. In Allen, J., Fikes, R., & Sandewall, E. (Eds.), *Proc. of the Intl. Conference on Principles of Knowledge Representation and Planning*, (pp. 473–484). Morgan Kaufmann publishers Inc.: San Mateo, CA, USA.

Russell, J. & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, *11*(11), 273–294.

Russell, J. A. & Feldmann Barrett, L. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, *76*(5), 805–819.

Scherer, K. R. (1984). On the Nature and Function of Emotion: A Component Process Approach. In K. Scherer & P. Ekman (Eds.), *Approaches to emotion.* (pp. 293–317). N.J.: Lawrence Erlbaum.

Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal Processes in Emotion* chapter 5. Oxford University Press.

Scherer, K. R. (2005). Unconscious processes in emotion: The bulk of the iceberg. In P. Niedenthal, L. Feldman Barrett, & P. Winkielman (Eds.), *The unconscious in emotion* chapter 13. New York: Guilford Press.

Sloman, A., Chrisley, R., & Scheutz, M. (2005). The architectural basis of affective states and processes. In *Who needs emotions?* Oxford University Press.

Wundt, W. (1922/1863). *Vorlesung über die Menschen- und Tierseele*. Leipzig: Voss Verlag.