CHAPTER 5

A Framework for Emotions and Dispositions in Man-Companion Interaction

Harald C. Traue, Frank Ohl, André Brechmann, Friedhelm Schwenker, Henrik Kessler, Kerstin Limbrecht, Holger Hoffmann, Stefan Scherer, Michael Kotzyba, Andreas Scheck and Steffen Walter

This research was supported by grants from the Transregional Collaborative Research Center SFB/TRR 62 Companion Technology for Cognitive Technical Systems funded by the German Research Foundation (DFG). Translation of the German version by Ute von Wietersheim.

Pleasant company alone makes this life tolerable. *Spanish.*

1. What are Companions as Technical-Cognitive Systems?

Digital companions are embodied conversational agents (ECA). They communicate in natural spoken language and realize advanced and natural man-machine interactions. It is the main goal of such companions to provide not only functionality but also empathetic responding to the user's needs. In terms of etymology, the English term 'companion'[1]

[1] Webster's New World Dictionary (1970) defines companion this way:
A person associated with another person
A person employed to live or travel with another person and
A member of the lowest rank order of knighthood

means *fellow, mate, friend,* or *partner*. It originates from late Latin and literally means *companies,* i.e. *with bread*, an individual close to us, which is able to give us something (in this case: bread). It is unclear whether it is a translation from the German word *gahlaiba*, which mutated in the German language to *hlaib* and finally *Leib* (Eng.: loaf). Companions are individuals close to us *with bread* that replaced the old English word *gefera*, the travel companion, which was derived at from *faran*, and in German to the word *fahren* (drive) and finally to the word *Gefährte* (which translates into companion, closing the circle between traveling together and being able to supply with food). What is also noteworthy is a cross connection to Arabic, because the word *Faran* is a male name, which means baker.

The etymological origin is relevant because current research not only focuses on technical realization problems, but the nature of possible relationships between humans and digital companions is under debate. This was especially expressed in the compilation of contributions titled "Close Engagements with Artificial Companions," which was published by Yorick Wilks in 2010 as a result of a comprehensive seminar at the Oxford Internet Institute. This compilation investigates the topic of cognitive-technical intelligence as a constantly available, selfless and helpful "software agent" from different perspectives as a future vision more intensively and with more facets than ever before. Aside from presenting already existing prototypes, work platforms, and application areas, this discussion gave a lot of room to ethical, philosophical, social and psychological issues because all experts taking part in the discussion believe that in just a few years companion technologies will have enormous communication capabilities. These communication capabilities will allow for human-companion interactions in many areas of life such as at work, in daily life, with regard to health maintenance, mobility, and social networking through highly selective information flows, which exceed the capabilities of currently available assistance systems, humanoid robots, or entertainment technology. This vision emphasizes the need for a practical theory on companion features that must take into account the psychological and social capabilities of cognitive-technical companions with respect to their human users. Companion technologies will make various sources of information available (for example, from the internet) for the interaction between humans and technical systems. Human-companion interactions will not be identical with human-human interactions, but they will probably be very similar. This similarity is maybe not due to the humanoid design of the companion, but the structural similarities of communication as well as information transfer/processing. When asked to compare and contrast positive and negative experiences with technical devices (human-machine interaction, HMI) or with other human beings (human-human interaction, HHI), and whether the emotional content of such experiences can be analyzed, there are significant similarities, but also a few differences because, with regard to HMI, feelings such as shame would be very rare, but do play a role in interpersonal relationships. In addition, negative emotions showed more variety in HHI (Walter et al., 2013).

Table 1. Fundamental differences between humans and cognitive-technical companions.

Characteristics	Human Companion	Cognitive-Technical Companion
Determination	Unsure	Determined
Materiality	Organismic, not deterministic	Technical and algorithmic
Availability	Depending on will and activation	At will
Autonomy, personality, and awareness	Yes	No
Emotionality	Subjectively experiencing, socially expressive and emphatic, embodiment	Sensory recognizant and expressive (Avatars)
Communication and Ability to talk	Potentially very comprehensive, multi-modal, natural speech	Very limited, multi-modal
Needs, motivation	Varied, psychobiological	Technical energy supply
Sensitivity	Mental and physical	Device-related

Table 1 lists some of the fundamental differences, which particularly refer to technical and biological characteristics. From the view of an interacting user, however, not all of the differences are relevant. What are most important are the corporeality of a human and its ability to communicate naturally, and the ability of the cognitive-technical system to interpret the natural speech of the users in a semantically correct manner. If the user can use natural speech to interact with the companion, albeit in a limited manner, the character differences become less important and the relationship that a human user develops with an object will depend on the emphatic capabilities or a technological system in responding to human emotions (valence, moods, and discrete emotions) and dispositions (motives, action tendencies, and personality). The user does not approach a new technical system with companion characteristics as an unknown entity, but will transfer his/her "inner world" of earlier effective experiences to the new situation, according to Kernberg (1992). From the self-psychology perspective (Kohut, 1987), it must furthermore be assumed that the symbolic and emotional inner representation of objects fulfills a function, i.e. to maintain and improve the functionality of the individual (for example, as an enhancement). Turtle (2010, p. 5) illustrates this assumption with the help of a student, who commented that she would love to replace her "real" boyfriend for a social robot if the robot were nice to her: "I need the feeling of civility in the house and I don't want to be alone … If the robot could provide a civil environment, I would be happy to help produce the illusion that there is somebody really with me."

A good companion makes a heaven out of hell. German.

2. From the Assistance to the Companion System: A Qualitative Leap!

Nobody would currently expect a navigation system to be able to respond to the frustrated undertones we use to respond to the repeated instruction, "*When possible, make a U-turn*" when traffic on the opposite side of the highway has come to a standstill. A human passenger would not be forgiven for saying that because he is aware of the situation and the fact that it is simply not possible to do a U-turn. Furthermore, we would expect that navigation system reacts with empathy to the driver's emotional response caused by the repeated insistence "*When possible, make a U-turn*". The continuous and senseless repetition of the instruction can cause anger in this case. Even more, a lower frustration tolerance as a personal characteristic can intensify this. A traffic jam can also create emotionality. Consequently, there are three possibilities: 1. The emotionality is the result of the interaction between a human and a companion, 2. The emotionality is the result of an external situation or 3. The emotionality is the result of both.

What would turn such an inadequate assistance system into a companion system? It would have to say to the driver, for example: "*Why are you not turning?*" and when the answer is: "*I can't. There is a traffic jam*" it would have to respond by saying: "*Okay, let me try to find another route, stay in this lane.*" Such a companion would, at that moment, be more communicative and competent than the driver, if it has information about the traffic jam, and it would be empathetic. The automotive industry is on its way there. It presented at the CeBIT 2011 the "connected car", an upgraded electric Smart, which allows the driver to activate many functions using voice commands. Furthermore, it makes use of external data sources and services over the Internet. For example, the driver can order movie tickets. Ford has presented its embedded *Sync* system, which, in case of an accident, automatically generates an emergency call with location information and informs the driver of the vehicle involved in the accident that help is on the way (Asendorpf, 2011).

Biundo and Wendemuth (2010, also refer to the research request by SFB-TRR 62[2]) describe companion systems as cognitive-technical systems, whose functionality is completely adapted to the individuality of its user. Companion systems are personalized in respect to the user's abilities, preferences, requirements and current needs, and reflect the user's situation and emotional state. They are always available, cooperative and trustworthy, and interact with their users as competent and cooperative service partners. The functionality named in the work definition is not further explained in detail, but it lists, as a central assistance function of a companion system, planning and decision-making systems with which the user and system are equally confronted. The assistance function is a decision-making support function which provides the user with options and the respective reasons that the user can then accept or reject. Companion technologies are therefore not only intended to be an improved interface. Furthermore, they should make the functionalities of the technical systems individually available, but also make new, complex application domains feasible. The following are often cited as applications or domains:

- Assistant for technical devices
- Household and telecommunication devices
- Entertainment electronics
- Ticket dispensing machines
- Medical assistance systems
- Telemedicine
- Organizational assistants
- Health prevention
- Support systems for patients in rehabilitation clinics
- Support systems for individuals with limited cognitive abilities and much more.

In all these application domains, planning and decision-making processes play an important role. In some examples such as telemedicine or organizational assistance, they have top priority with regard to the explicit functionality (Biundo and Wendemuth, 2012). Immediate emotions have an impact on information processing, planning and decision making (Loewenstein and Lerner, 2003). Emotions are therefore considered not only as a subjectively experienced emotional

[2] www.sfb-trr-62.de

by-product, but as an action-managing affect factor of companion functionality. Sloman (2010) argues that the complexity of companion features as such should be discussed, but so should the quality and complexity of the requirements. His starting point is the presentation of an illustration of possible interaction flows (world knowledge), with which a companion can compare the respectively current behavior and therefore knows "*what they have done, what they could do, what they should not do, why they should not do it, what the consequences of actions will be, what further options could arise if a possible action were performed, how all this relates to what another individual could or should do, and can also communicate some of this to other individuals*" (p. 180). This would create some sort of situational awareness of the companion, not necessarily a consciousness (that would relate to a subjective experience). At this point, it is apparent that the goals that companion technologies have set for themselves can only be achieved with patient scientific work, because the euphoria about the thinkable and desirable system features of companions should not lead to the wrong conclusion that the necessary formal descriptions of these system features have yet been solved and if, in a rudimentary fashion, only relates to very special, mostly simple cases. In this context, Sloman (2010) initially views two very narrowly defined "target functions" that will become relevant for companions in the near future, and that can be assigned to companion technologies, i.e. "engaging function," which mostly refer to the quality of the interactions that entertain, draw attention, are fun or are just interesting and "enabling functions," which support users with regard to their goals, motives and intentions. The latter can help to solve various everyday problems, since they provide information, teach, organize the user's social and physical environment and enable the user to participate in society. Sub-functions of these target functions are currently offered by existing companion-like assistance systems, which are therefore very helpful as an inspiration for our visions of future technologies.

Beverly Park Woolf from the Department of Computer Science at the University of Massachusetts leaves no doubt: "*If computers are to interact naturally with humans, they must express social competencies and recognize human emotion*." Using the example of tutorial companion systems (CS), she shows that the sensory capturing of dispositions such as boredom, interest, and frustration by companions makes the tutoring functions of learning supports significantly more effective, increases motivation, and reduces adverse emotional states such as frustration, anger, or fear (Woolf, 2010, p. 5). The companions that Woolf refers to as social tutors capture emotional and dispositional responses during the learning process by measuring posture, movements, grip strength, physiological agitation, and facial movements through sensors in the chair, the monitor, the mouse, and the skin. A relatively simple model is used, in which the values of four parameters can be allocated to the following four dispositions (here referred to as emotions) with a precision ranging from 78% to 87.5%: boredom, flow, interest, and frustration. Empirical tests have shown that by using emotionally adequate responses (50 variations of support and encouragement, to keep going and trying hard), students work significantly longer on frustrating and difficult tasks and that their stress level is decreased. The tutorial systems not only respond verbally, but are also able to show an emotional response to the users in the form of avatars, or support the interaction with an interested facial expression as well as positive gestures. The empirical research has shown significant gender differences: Female test subjects benefited more from the emotion-based responses of the digital tutor than male test subjects.

Simply because of the fact that language is the most important form of communication for human beings shows that companion systems should also be able to communicate verbally. Nass and Brave (2005), however, impressively documented which possible consequences must be considered when technical systems are equipped with verbal interaction capabilities. Verbal technical systems could potentially be associated with certain social competencies or personality traits that could have a conscious or unconscious influence on the user's interaction with a technical system. If the expectations are not fulfilled, the system might be less accepted and not trusted. Notwithstanding these risks, however, verbal communication also offers the unique opportunity to increase trust and acceptance not only by providing information in an effective and natural manner, but also by using voice modulation (prosody) capabilities. Humans are probably more aware of what is being said in a conversation, but how something is said is often just as important, especially in social interactions. In the tutorial area, the prosody tool is used to positively influence the learning situation. Prosody can, for example, support the motivating character of comments and help in learning situations to increase perseverance and avoid errors. Consequently, this tool also seems suitable for interactions with technical systems to increase effectiveness. While there have been numerous studies in recent decades about the production and processing of prosody in human-human interactions (Frick, 1985; Scherer, 2003b; Baum and Pell, 1999; Friederici and Alter, 2004), this topic has not yet been investigated intensively with regard to the human-machine interaction area. Until now, the recognition and classification of human prosody by the technical system have been the main topic of the research that has been done in this regard (Cowie and Cornelius, 2003; Schuller et al., 2003). The question of how the use of prosody by a technical system can positively influence human users is still fairly unanswered. A recently published study was able to show that praising and reproaching prosody instead of neutral comments from a system (such as "*correct*", "*incorrect*", "*yes*", "*correct*", "*no*", "*incorrect*"), in response to answers selected by test individuals in a learning situation, can lead to significantly higher rates of learning (Woolf, 2010). In addition, this study showed that the use of a synthetically generated version of the spoken comments (compared to prosodically neutral, naturally spoken comments) led to a significantly worse learning performance. This was all the more remarkable because the information about the accuracy of the test person's answer was also contained in the synthetically generated response and therefore the decisive information to solve the task. These findings emphasize the risks of using verbal communication, especially by companion systems, because the scope of such systems' communication capabilities must be so complex that a naturally spoken vocabulary would not be practical in contrast with a mere navigation system and its limited vocabulary.

He who has a companion has a master. French.

3. Relationship between Humans and Companions

Why should a human being surround himself with a companion that is more than just a passive assistance system because it reacts emphatically to his or her emotions, needs, and motives? Two perspectives are important here: First, users want to reach a goal, solve a problem, or improve their capabilities. Taking a mountain hiker as an example, the companion could be used to reach a certain objective or avoid bad weather. The interaction with a companion supports users with its specific problem-solving competencies and ability to carry on a dialogue. The companion system takes into account the user's emotional and motivational state as well as his or her cognitive abilities to solve a problem. Therefore, if the fact that bad weather is approaching creates concern, the companion would urge the user to avoid the weather and if that is not possible, it would help prepare for surviving bad weather in the open. In order to do so, the companion must have specific capabilities and must be able to dialogue. In this case, the emotions, motives and intentions are created by the context in which the need for using the system originated. Second, usage is the result of a need for contact, entertainment and activity. To be able to do so, the companion system also requires specific capabilities, but is not activated in the interest of solving a problem. A mountain hiker could also use the companion for entertainment or relaxation purposes when taking a break. Emotionality and motivation change in this case in the dialogue with the companion system itself and not necessarily in the achievement of goals.

If companion systems are able to perform certain technical functions, independently perceive their physical and social environment with the help of their cognitive capabilities, map this information in internal models, and draw conclusions from this information and embed it in internal plans and objectives so that they can subsequently communicate with their users, for example to align human intentions in certain situations with factual requirements, then companion systems are capable of doing things that otherwise only humans can do in an interaction: they can be a friend giving advice, a guide, a therapist, a coach, an expert, or a teacher. It is also feasible that the companion can be used to support an inner dialogue. Humans often use such inner dialogues to look for support in ambivalent or critical situations, to substitute something, to explore possibilities, to bond, to improve themselves, to gain insights, or to self-guide themselves. In such a function, the companion system could use the user's voice (Puchalska-Wasyl, 2007).

To the extent to which the companion's functions relate to its empathy and adaptability and to the extent its communication behavior is geared toward the user's individuality, users will develop a relationship, feelings and a bond with the companion, in which there is an I, a YOU, and a social environment. Users will also form a model of the companion that reflexively includes assumptions about the user model in the companion. Users have preconceived notions about the companion's characteristics and will continue to dynamically develop these over the course of the interaction (also refer to Figure 1).

The quality of such a relationship between human and companion depends on different factors: Prior experience (priming as described by James et al. (2000)) and attribution of the companion's behavior by the user (Bierhoff, 2011) and the projection and transfer of the user's conscious and unconscious wishes and expectations to a given companion system. These factors describe cognitive filters that influence the processing of information in the HCI, but are not created per se in the context of the HCI. They describe earlier experiences, expectations and personal characteristics of the users, if these are relevant for the HCI. Turtle (2010) describes this facilitated projection process, when not just the function but also the design is human or animal-like (even without any special cognitive functions): When

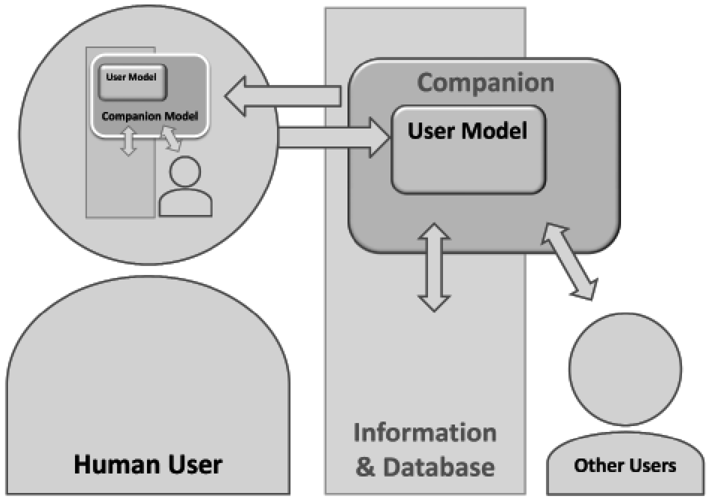

Figure 1. Reflexive nature of the user model in the companion system and the working model of the companion's user.

robots make eye contact, recognize faces, or mirror human gestures, they push our Darwinian buttons, exhibiting the kinds of behavior people associate with sentience, intentions, and emotions. Once people see robots as creatures, people feel a desire to nurture them. With this feeling comes the fantasy of reciprocation: as we begin to care for robots, we want them to care about us.... Eleven-year-old Fara reacts to a play session with Cog, a humanoid robot at MIT by stating "it's like something that's part of you, you know, something you love, kind of like another person, like a baby." (p. 4)

Keep company with good men and good men you'll learn to be. Chinese.

4. Process Component Model of Moods, Core Affect, Emotions, and Dispositions

4.1. Emotional and dispositional behavior components

The user's mental states that are relevant for companion technologies are summarized under the term emotions and dispositions. These refer to the totality of moods, emotions, motives, action tendencies, and personality. The special companion feature, i.e. the ability to empathically recognize mental states and be able to adjust its own technical functions to the human user, is intended to increase the acceptance of technical systems for certain functions that are useful to humans, and thus make them continuously available to the human user via an interaction cycle. This is to prevent reactance as the result of insufficient empathy.[3]

Emotions and dispositions[4] comprise all psychobiological states that, at varying degrees and complexities, influence the dialogue between a human user and the companion as well as its functional use: Newness and valence, core affect, discrete emotions, moods, motives, action tendencies, and personality. Table 2 hypothetically describes several criteria for the different emotions and dispositions. The time dynamics for the situation assessment of valence (positive vs. negative) is very fast, probably in the range of 200 ms. By contrast, a user's personality only changes very gradually or not at all over months or years. The influence on behavior management probably acts according to a U-function, because the fast assessment of the valence has a strong influence on behavior as well as personality. Moods and action tendencies have less influence within this hypothetical model. The feedback strength

Table 2. Hypothetic illustration of emotions and dispositions as well as characteristic developments for dynamics, the strength of behavior management, feedback strength, and complexity

Characteristics Emotions and Dispositions	Dynamics	Impact on Behavior Control	Feedback Strength	Complexity and Operationalization
Newness and Valence	ms–s	Very high	Very high	Bi-modal and simple
Core Affect and Discrete Emotions	s–min	High	Very high	Average and difficult
Moods	min–days	Average	High	Average and difficult
Action Tendencies	min–hours	Average	High	High and difficult
Motives	hours–months	High	Average	High and simple
Variable and Stable Personality Traits	months–years	High	Low	High and simple

[3] However, it is quite possible to imagine situations in which a user expects decisive interventions from the companion technology, or, for example, when the companion technology is expected to avert danger. The prerequisite is an initiated dialogue between human and companion, during which the realization of companion characteristics is agreed upon (Bryson, 2010).

[4] Disposition (v. lat.: disposition = distribution, allocation, structure, listing, plan). In psychology, the term is used within the meaning of a readiness for (usually pathological) reactions, but is mostly used as a common expression. According to the general definition, which suits the term disposition as it is used within the context of man-companion relationships is "the organized totality of the individual's psychophysiological tendencies to react in a certain way" Chaplin, J.P. (1968, 1975) Dictionary of Psychology. New York: Dell Publishing Co.

describes the immediate effect of a disposition in interaction sequences. A necessary condition for immediate impact is the dynamic of emotions and dispositions, since feedback is only possible if there is change. The strength of the feedback and the impact on behavior control are correlated. Due to the complexity of emotional responding, some emotions are difficult to measure. Personality is also complex, but can easily be measured with self report scales and there is no need to capture personality over time during the interaction. That is also the case for the motivational structure. Action tendencies, discrete emotions and moods must be dynamically captured. So far, that has not been sufficiently achieved. Valence is bi-modal or a one-dimensional value and dynamically recordable.

Emotions and dispositions are the result of the processing of information about emotion-relevant stimuli and their unconscious and conscious cognitive assessments. Dispositions comprise the willingness to respond to emotional and non-emotional stimuli. Emotions and dispositions have an influence on each other. The general consensus is that emotions are composed of several components (Frijda, 1988; Scherer, 2001; Traue and Kessler, 2003; Traue et al., 2005; Frijda, 2007):

- Subjective experience (feeling, mostly semantically codable)
- Cognitive assessment of inner and/or outer stimuli (appraisal)
- Expressiveness of facial movements, gestures, and the body as a whole
- Psychobiological, neuronal, and endocrine activation
- Cognitive drafting of action tendencies and actions

The sensory groups that capture the respective behavior of an individual are also structured according to these components. Each of the components has its own chronological dynamic. This dynamic and the pattern of the sensory parameters result in a clear allocation of emotional processes in humans. Emotions are subjective experiences that, in different situations, are perceived similarly by different individuals. Emotions can also be understood as flexible adjustments between an individual reaction and situation, which lead to action tendencies and facilitate intra-individual and inter-individual interaction regulations (Traue and Kessler, 2003).

With regard to the description of emotions, a distinction can be made between the structural and functional views. The structural perspective describes the inner relationship of emotional components for the temporal processing. For the objective of differentiating between emotions with pattern recognition processes, the structuralism position is particularly suitable (Witherington and Crichton, 2007). From the functional perspective, the emotion components serve different goals. For example, facial expressions serve to communicate the emotions in the social environment, cognitions serve to evaluate stimuli and serve to plan behavioral activity, while physiological reactions, among others, regulate the energy budget, and finally subjective experience serves the conscious awareness of emotions. Emotions are behavioral units, whose components belong to each other from a structural perspective, that develop through an interaction between emotional stimuli and the individual over time and that lead to a process: "Whereas the functionalist approach focuses principally on the nature of emotion, the dynamic systems approach focuses principally on the nature of emotional development on the process by which emotions emerge in real-time contexts and undergo change across developmental time" (Witherington and Critchton, 2007, p. 629).

Moods are potentially long-term emotional states that may affect the quality of the individual experience, but are less intensive. Discrete emotions emerge from mood states. Frijda (1988) sees in the blocking of actions and the triggering of action tendencies a main component of emotions: "Individuals experience the urge to come closer or to turn away, to start screaming, or to sing and move; some just want to withdraw and do nothing, to no longer have any interest or to lose control" (1988, p. 351).

Certain emotions can be allocated to the initiation of action tendencies: Positive emotions activate people to approach other people and objects. In the form of desire, it is a strong, contact-promoting emotion. In contrast, fear triggers avoidance, but also the need for protection or help. Anger leads an individual to turn to someone because it creates the mental energy for coping with or even eliminating the issue. Sadness is an approach to loss. This emotion serves the (imagined) existence of a lost object. Contempt is the socially expressive avoidance, but also fear leads to avoidance. Surprise creates attention and interest. More complex and secondary discrete emotions, such as embarrassment or shame, lead to less clear action tendencies because they usually depend on complex appraisals. The structural similarity of secondary emotions such as guilt, shame, embarrassment, pride, self-confidence, honor or jealously consists in their dependence on the relationship between the ideal and the real self and therefore on complex, conscious, cognitive appraisals.

Emotional components of an emotional event in an individual do not follow the same time course. Unconscious to the individual, the components may follow their own dynamic process: Processes of the central nervous system may only take fractions of seconds and lead to a first, quick appraisal, whereas more complex, cognitive appraisals and subsequent behavior like an approach or facial expression may last several minutes. The stimulation of any given component and its progression are different and depend on each other. The reciprocal dependence of the emotional components and their different dynamics create recursive effects between the consecutive early and late emotional appraisals, the emotional reaction, and the action tendencies. The consecutive elements influence, in the form of feedback, the emotional evolution with regard to the chronological progression and the emotional quality (Colombetti, 2009).

4.2 Cognitions: Sequential check theory

The currently influential component process models (CPMs) describe the five emotional components listed in Section 4.1 and their high interdependency (Scherer, 2001). These models are based on the assumption that perceived emotional stimuli of the social and physical environment require cognitive appraisal, and try to further specify the connection between a specific stimulus and the pattern of the resulting physiological, expressive and motivational change in order to map an integration of the entire emotional process. The cognitive component in which the appraisal or assessment of inner and/or outer stimuli takes place plays an important role.

According to Leventhal and Scherer (1987), cognitions check sequentially every inner and outer emotional stimulus. This cognitive process, described as Sequential Check Theory (SCT), should be understood as a part of the dynamic CPM, where the cognitive appraisal takes place. The sequential check theory tries to explain the different emotional states of an individual as the result of a specific stimulus evaluation check and makes predictions about subsequent response patterns in the individual organic sub-systems. The stimulus evaluation is divided into four main steps that are necessary for an adaptive response to an emotional stimulus: Relevance, Implication, Coping Potential, and Normative Significance. The assessment is subjective and does not have to match the objective characteristics of a given situation. Furthermore, it should be noted that the assessments take place both subconsciously and consciously (see Figure 2).

The relevance assessment in the *first step* determines how important the event is for the individual. Both the external and the internal environments are monitored with regard to the occurrence of potentially emotional stimuli that require allocation of attention, further information processing, and possibly adaptive reactions. The

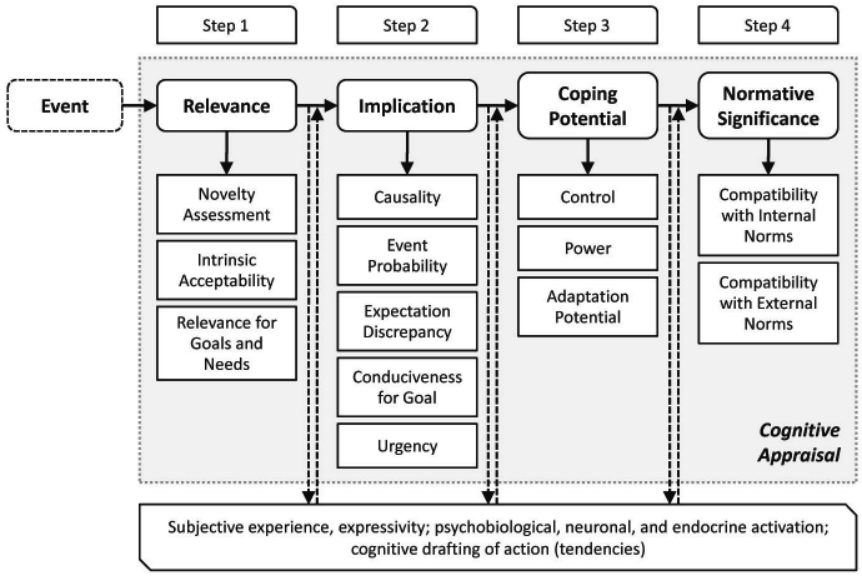

Figure 2. Stepwise Cognitive Appraisals with checks of relevance, implication, coping potential, and normative significance as a result of emotional events according to in the Sequential Check Theory.

relevance detection is divided into an appraisal of novelty (within the meaning of sudden onset and familiarity of the events), intrinsic acceptability, and relevance for goals and needs.

The *second step* comprises the assessment of the implication, and deals with an estimation of consequences and future developments, whereby the causes and the consequences of the event are considered. This process is divided into the following five assessment dimensions: Causal attribution, meaning who did what and why?; the probability with which certain consequences are expected; the discrepancy or consistency with expectations that the individual has regarding the situation (created by the event) at this point in time; the positive or negative effect it has on the individual's goals and needs; and the urgency that depends on the importance of the events for the individual as well as chronological reaction contingencies.

Step three is responsible for evaluating the coping potential, which is assessed in order to change an event and its consequences or to successfully adapt to unchangeable events. The control over the event or its consequences, the power (of the individual to be able to align the results of the event with his/her own interests, when he/she has some measure of control) and the adjustment potential to the event play an important role. The *fourth step* checks the normative significance, in which the individual assesses how the majority of other group members will assess his/her actions. A distinction is made here between internal and external standards. Internal standard checks determine the extent to which an action corresponds, for example, to one's personal self-image. In the external standard check, the compatibility of an action with perceived standards is reviewed.

Leventhal and Scherer (1987) postulated that the evaluation steps follow a fixed sequential order. This sequence supposition is explained from system economics and logical dependencies. However, the behavioral and subjective results of every step influence the next cognitive evaluations which, in turn, influence the evaluation process in the individual steps (shown by dashed arrows in Figure 2).

Since it can be shown that moods, emotions, motives, individual personality differences and even cultural values and group pressure greatly influence the result of cognitive evaluations, these determining factors must also be taken into account by the modeling of user emotions in companion sytems.

Table 3. Emotion processing system on the sensory-motor, schematic, and the conceptual level, according to Leventhal and Scherer (1987, p. 17).

	Novelty	**Pleasantness**	**Goal/Need Conductiveness**	**Coping Potential**	**Norm/ Self-Compatibility**
Sensory-motor Level	Sudden, intensive stimulation	Intrinsic preferences/ aversions	Basic needs	Available strength	Emphatic adaptation
Schematic Level	Familiarity: Pattern comparison	Learned preferences/ aversions	Acquired needs and motives	Body schemata	Self/Social schemata
Conceptual Level	Expectations: Cause/ Effect, Probability	Remembered, anticipated or deducted positive-negative assessments	Conscious goals and plans	Problem-solving ability	Self-ideal, moral evaluation

Leventhal and Scherer (1987) present an emotion-processing system, in which the evaluation process takes place on three different levels: The sensory-motor, the schematic, and the conceptual level (see Table 3). These three levels could also be aligned with the evaluation processes of the sequential check of emotional stimuli. On the sensory-motor level, the assessment of events takes place mainly on a subconscious level based on intrinsic functions and reflexes. On the schematic level, social, individually learned patterns are used to evaluate the event, which are mostly automatic and therefore rather subconscious. A situational similarity recognized by the companion could help here, in case of a conflict between the situation and the user (motives), to adequately solve it. The sensory-motor and schematic level of cognitive processing are assessable by extended measurement of psychobiological sensors of the motor, autonomic and central nervous system and by automated classification procedures. The last conceptual level now allows information stored in the memory to be intentionally used for the evaluation in a reflexive, conceptual-symbolic process in humans. On this level, it is appropriate to have the companion system run semantic analyses to anticipate the user's emotional and dispositional state.

The cognition-theoretical formulations (Leventhal and Scherrer, 1987) differentiate appraisals by complexity, cognitive content and the level of control between automated and intellectually derived. The automated appraisals are only cognitive in as much as all higher functions of the brain can be referred to as cognitive. These appraisals are not necessarily conscious; however, its result could be an emotional, subjective experience. Even evaluations of goal achievement or impairment can run in an automated and unconscious manner, but will create a subjective feeling. It should be mentioned that appraisal theories, which consider cognitive stimuli evaluation as the main causal factor for emotions, are not generally accepted. According to Frijda (Frijda and Zeelenberg; 2001; Frijda, 2007), evaluations are the result of a monitoring process of actions and intentions. He believes that such evaluations do not cause any emotions, but rather, are cognitive accessory phenomena of the situation or the action, but not causal for emotions. Whether it is accessory phenomena or causal triggers, it can only be answered for human-companion interactions if researching the subjectively felt emotions (feelings) and the other components of the emotional behavior with regard to their chronological development.

4.3 Dispositions: Motives, action tendencies, and various personality traits

4.3.1 Motives

A user's needs, which are relevant for the interaction between humans and companions, can be categorized by deficits and the needs for growth (deficit needs and growth needs within the meaning of Maslow (1943)). The basic needs on the highest level characterize the need for respect by others, social acceptance, and self-realization (also refer to personal growth as per Bandura (2001) and Aubrey (2010)). These include[5]:

- Improving self-awareness and self-knowledge
- Building or renewing identity
- Developing strengths or talents/potential
- Enhancing lifestyle or the quality of life
- Improving health
- Enhancing personal autonomy
- Improving social abilities

Such needs can govern the intensive use of complex and socially competent companions. With regard to deficit needs, companion technologies can compensate for a lack of safety and order. Maslow (1943) also lists belonging and **attention as important deficit needs.** Schuler and Prochaska (2000) distinguish between differentiated primary motives, social motives (dominance, competition and status orientation) and performance motives (e.g. commitment, the willingness to put forth effort, and persistence). There is a respective measuring instrument for this empirical motivation model (*Leistungsmotivationsinventar [performance motivation inventory], LMI by* Schuler and Prochaska, 2000). In the available test version, the scales included in the LMI refer to self-assessments and therefore allow documentation of the cognitive context that a user brings with him/her when engaging in a dialog with a companion. At the same time, the scales can also be used to describe behavior characteristics in the HCI because they describe action tendencies. Two scales from the total of 17 personality variables of the performance motivation are particularly relevant for tutorial applications. Persistence is defined as perseverance and the use of energy with which tasks are handled. Confidence in success describes the optimistic attitude toward difficult tasks that the abilities, skills, and knowledge will successfully lead to the desired goal.

4.3.2 Action Tendencies

Action tendencies are the result of imbalances of emotions and the behavioral consequences of motives. For a companion technology, it is therefore important to recognize changes of emotional components and the subjective experience of the balance. Vigilance, selective attention, approach/avoidance, interest, frustration and conflict/ambivalence are companion-relevant action tendencies.

[5] http://en.wikipedia.org/wiki/Personal_development

Vigilance refers to the sustained attention (German: Wachheit) during a certain period of time. Vigilance is the requirement for conscious information processing. It correlates in a causal-functional fashion with the stimulation of the central nervous system. The two poles on the vigilance continuum are high activation, e.g. extreme stress or startle and slow wave sleep. *Vigilance* refers also to the ability to respond to accidental, low-threshold, and seldom events in a meaningful manner. The vigilance stages can be measured continuously with the electroencephalogram.

Selective attention is the limited ability to simultaneously pay attention to multiple stimuli or sensory modalities. The reason for this limited ability is the assumption of limited information-processing capacity of an individual. The selective attention (also concentration) describes the focused attention on certain stimuli, mostly provided within the context of a task, while other stimuli can be ignored. Eye tracking would be an appropriate measure.

Avoidance is an action tendency to withdraw from a situation or action. It is triggered by a (conscious or unconscious) assessment of the situation as unpleasant, dangerous, or threatening. Also, a threat to one's self-worth or the anticipation of effort can trigger avoidance. Avoidance of a behavior triggered by (anticipated or imaginary) ideas can protect from unpleasant states, but also prevents new and positive experiences. Avoidance behavior is behavior that is learned through a combination of traditional and operant conditioning or by learning from role models. The self-reinforcement of avoidance behavior by negative reinforcement turns avoidance behavior into a stable behavior pattern.

Interest is a form of selective attention, which is referred to as the cognitive participation and attention to certain topics, tasks (for example, the reading of information) and content areas. It classifies the interests of a person for certain things (e.g. professional interests, hobbies, or political interests). Modern interest theories and research approaches (Krapp, 2002) describe a person-object concept, in which the degree of interest is defined by the subjective appreciation of an object area. This term is particularly relevant for tutorial systems because interest is defined there as the emotional, motivational, and cognitive interaction between a person and his/her object areas. Lack of interest can therefore be described as distraction, lack of selective attention, etc.

Frustration is created when a person is prevented from reaching a goal because of real (external) or imagined (internal) reasons. The intensity of the frustration depends on the attractiveness of the goal and the motivation to reach the goal. An emotional response to frustration may be anger and regression (helplessness). A subsequent action tendency may be either approach or avoidance.

Conflict and ambivalence refer, in the man-companion interaction context, to competing motives or action tendencies that can be triggered by ambiguous or several stimuli. The desired objectives or goals (real or imaginary) have either an appetence (the individual pays attention to it) or an aversion (the individual does not pay attention to it). Since an objective may trigger several appetences or several aversions or both, the result for two objectives/goals, according to Miller (1959), is as follows:

1. Appetence-appetence conflict: Both objectives/goals are considered positive,
2. Aversion-aversion conflict: Both objectives/goals are considered negative or
3. Appetence-aversion conflict (ambivalence conflict): Both objectives/goals are considered both, positive and negative.

Flow: If a dialogue behavior is mainly controlled by external stimuli (references, reward), inner involvement decreases. To maintain a difficult, task-related interaction with a technical system, the joy the activity brings should be the motivator. Such states are often referred to as flow (Keller et al., 2011).

4.3.3 Stable personality traits: Optimism, hardiness and a sense of coherence, NEO-FFI, emotion regulation, and attribution style

Personality traits describe individual differences that affect emotions and dispositions because they affect the perception of internal and external events and their cognitive and emotional processing. Processing introversion correlates with the intensity of the psychological stimulation during negative emotions, the tendency not to show emotions (suppression in ERQ) influences the reduction, the facial expression, as well as increased psychophysiological reactivity (Traue and Deighton, 2007). All cognition-related personality characteristics such as need for cognition, attribution style, etc. will impact stimuli and coping appraisals. The sociability scale of the NEO-FFI has a strong impact on the social action tendencies, and scales such as optimism, hardiness, coherence, etc. are important in coping with stressful interactions (Traue et al., 2005).

Optimism is referred to as the positive general belief that one has enough resources to cope with stress. It is not important that such "subjective optimism" is justified, but the mobilization of behaviors and cognitive patterns enables the respective individuals to cope with difficulties.

These assumptions are based on research regarding self-perception and personality traits that show that mild and permanently positive illusions about one's own person and overestimation of one's own control of situations has a positive impact on self-confidence and the manner in which challenges are handled (Maruta et al., 2002).

In life event research, these abilities to meet stressful situations with *resistance* are defined with the key word **hardiness**. Hardiness is understood as the cooperation of three attitude and behavior patterns: *Control, Challenge* and *Commitment.* This kind of ideal-typical person always trusts in his/her own abilities, even under difficult life circumstances. Such a person considers life a challenge, in which every change can also be considered an opportunity and is mostly without any ambivalence both in his/her private and professional life, has few doubts, and is usually very dedicated and motivated.

The term **coherence feeling** is used for a global orientation that expresses to what extent an individual has a generalized, lasting and dynamic feeling of trust that his/her own inner and outer environment is predictable and that things will, in all likelihood, will develop in the manner that can be reasonably expected. Antonovsky (1987) proposes three components that relate to each other: *Comprehensibility* refers to the extent to which stimuli, events or developments can be perceived as structured, orderly and predicable. *Manageability* refers to the extent to which an individual perceives appropriate personal and social resources that can help cope with internal and external requirements. *Meaningfulness* finally refers to the extent to which an individual perceives his/her life as meaningful. In particular, the meaningfulness component puts the coherence feeling in a closer relation with emotional behavior because the assignment of situational meaning is a central emotional process.

The term **personality** refers to all mental characteristics of an individual that it shares with others or in which it differs from others. Widely accepted are five-factor models of personality: Extroversion, neuroticism, openness, conscientiousness and agreeability. These are described as stable, independent and fairly culture-stable factors. Extroversion is characterized by an outward-looking attitude. Individuals scoring high on the extroversion scale can be described as active, social, cheerful and/or talkative. Neuroticism, which is also referred to as emotional instability, describes the experience of and coping with negative emotions. Individuals scoring high on the neuroticism scale often experience fear, nervousness, stress, sadness, insecurity and embarrassment. The *openness* factor describes the degree to which an individual shows interest in and seeks new experiences. Individuals scoring high on this factor are often characterized as artistically inclined, imaginative, inquisitive and intellectual. *Conscientiousness* describes the degree of reliability, organization, deliberateness and efficiency an individual displays. The *agreeability* factor mainly describes to what extent an individual is altruistic. The more agreeable an individual is, the more empathetic, understanding and cooperative that individual can be described. The five factors can be measured with a standardized questionnaire, the NEO-FFI (Costa and McCrae, 2002).

The term **emotion regulation** refers to the ability to influence one's emotions in an active and targeted manner and not to interpret them as the consequence of another person's actions or the environment, which one cannot control. Emotion regulation consists of the following steps: the experience feeling must be detected, followed by a reflection about which response would be appropriate in order to avoid any reflexive or impulsive actions. Individuals with good emotion regulation show indications for mental diseases less often.

The **attribution style** defines which type of cause attribution an individual performs in order to explain his or her own behavior or the behavior of others. Different researchers have proposed different dimensions describing the attribution. The most frequently used dimensions are the distinction between internal/external and stable/variable.

The **need for cognition** is a personality attribute that describes how often and how much an individual likes to think about a topic. Individuals scoring high for this attribute enjoy thinking intensively about various situations and topics. Opinions are formed by way of an intensive review of the arguments. An exchange of opinions may therefore lead to a stable change in opinion. Individuals scoring low for this attribute generally use peripheral attributes such as attractiveness, credibility, etc., but the quality of the arguments seems rather unimportant. When such individuals change their opinion, the status is unstable, which is why it seems to be much more difficult to predict such an individual's behavior than for an individual with a high need for cognition. This construct can be measured with a standardized questionnaire, the need for cognition scale (Bless et al., 1994), and will be addressed in further detail in the operationalization section.

4.5 How to embed the process model in a companion system

At the beginning of an emotional event, there is an exogenic or endogenic stimulus. The individual confronted with the event is in a predefined state. This state is a result of the social context, prior cognitive activities (priming), the motivational situation and personality (Garcia-Pieto and Scherer, 2006). It switches like a filter between exogenic and endogenic stimulation and that furthermore modulates the emotional response. Figure 3 shows the progression and the structural connections for a single emotional behavior sequence consisting of the stimulus and the response. Since the emotional response itself acts within seconds as an endogenic trigger stimulus, it is reflected as a response.

The entire emotional process including its detailed recursion is also shown in Figure 3. In this chart, the initially simple (primary appraisal) and later more complex (secondary appraisal) cognitive responses are shown. These cognitive appraisals relate to the emotional stimuli, the coping competency and personality characteristics that tend not to change (for example, expressive suppression). Significance or meaning (stimulus appraisal) is attributed in several steps: First, the newness factor is assessed by comparing the event with memories in the working memory without any further cognitive involvement. If an orientation reaction takes place, it is a stimulus that is perceived as new. The stimulus is then assessed as positive or negative, depending on its relevance for the individual (preferences need no interferences, Zajonc, 1980). This primary appraisal or relevance detection process (Scherer, 2001) triggers the actual emotional response with behavioral, cognitive and psychobiological components. These may be primary emotions or, if less discrete, a shifting of the core affect in the three-dimensional space spanned by the dimensions of valence, arousal and dominance.

The emotional response is able to interrupt the current process. A strong fear response, for example, leads to a freezing of all movement. This emotional response is experienced as subjectively gestalt-like, whereby individual components most certainly can be perceived in a differentiated manner. Once formulated, the emotional response and the triggering event are subjected to an iterative, cognitive process for implications, coping potential, and ultimately, normative standards. The secondary appraisal is complex because it evaluates the individual emotional response by interoception (awareness of bodily responses) of the physiological activation, uses experience-based memories for coping strategies and because it must evaluate the necessity of adherence to norms. During this appraisal step, the individual checks whether the emotional stimulus is conducive or an obstacle to achieving a goal. The result may be fear and anger as a response to the interruption of a planned chain of actions. If the stimulus is conducive to reaching the goal, the individual might experience satisfaction or joy. In a last step, the individual's coping capabilities regarding situations are reviewed with regard to his or her goals and plans. The basis for this assessment is a causal attribution, i.e. the determination of what caused a certain stimulus. Without this causal attribution, it is often not possible to assess coping capabilities. If the individual cannot cope with the respective stimulus constellation without putting his or her important goals at risk, the result is anger or, in the event of habitual insufficiency, helplessness or depression. Finally, the relevance for the individual's self-image is processed. In an unfavorable case, it coincides with feelings of embarrassment, shame or guilt. Also included in this complex stimulus processing is information about the external stimulus, aspects of the self-image and especially social norms. In total (possibly after some back and forth iteration), this complex cognitive assessment leads to a determination of action tendencies and ultimately actions. In this process, the importance of the triggering stimulus may have changed.

The perceptions of the emotional response managed are not always identical and sometimes are even conflicting action tendencies. An emotional anger response may, for example, trigger action tendencies to show the anger in one's face and body language. The individual perceives this response and action tendency simultaneously, which can block his or her action tendency, depending on the social norm an individual has internalized (also refer to Traue and Deighton (2007)). These considerations lead to the recursion (Figure 3) of the various components (Scherer, 2003b).

The linear process character of an isolated emotional event (feed forward) starts from the stimulus and then proceeds via the primary appraisal and differentiation as well as the emotional response and the control of the action tendency and evaluation of behavior options

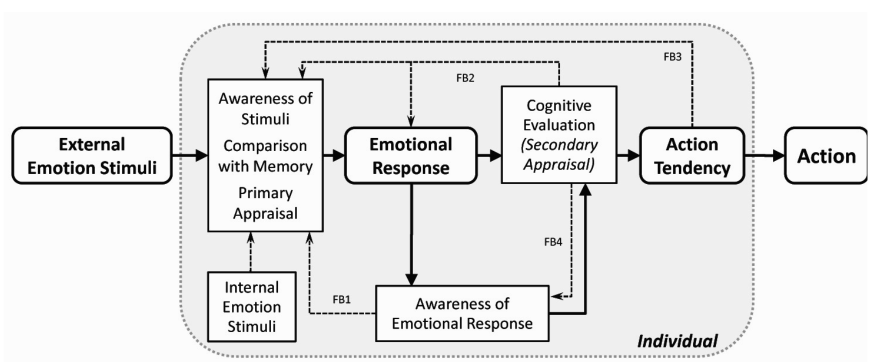

Figure 3. Recursive effects between emotion stimulus, emotion response, cognitive appraisal and action tendencies (adapted from Traue, 1998; Scherer, 2003b).

for the behavior. The first feedback (FB1) influences the primary stimulus appraisal by perceiving the response itself (e.g. the greater the anger, the more upsetting the situation). The second feedback (FB2) can influence the emotional response itself (e.g. keep cool, a little nervousness is okay), and the third feedback (FB3) affects the primary stimulus appraisal by assessing coping abilities or desirability (e.g. to make an omelet, you have to break a few eggs). Finally, the resulting action tendency may also alter the appraisal of the individual emotional response by feedback loop (FB4) (e.g. it annoys me, but I am not willing to do anything about it).

A man should take as companion one older than himself. African.

5. Modeling of Emotions and Disposition with COMPLEX

Based on a model used to simulate the interaction of artificial agents (SIMPLEX, Simulation of Personal Emotion Experience, Kessler et al., 2008), an expanded model was developed, which can map and formalize the interaction of a human user with a companion system (COMPLEX, Companions Personal Emotion Experience). Aside from the dynamic mapping of emotional and dispositional states within a technical-cognitive system, the emotional responses of users can be simulated and predicted in consideration of internal and external events, so that the functionality of the companion system can be enhanced in a meaningful manner.

External events (for example from the environment) are subjected to an individual appraisal process, based on the respective response to the event, in consideration of the available knowledge base (user or domain knowledge) and then individual goals can be determined. In addition to the appraisal of external events, internal events (such as psychophysiological parameters) can serve as input signals for the appraisal process. To customize the model, the values determined with such assessment processes are specifically modified on the basis of variable and stable personality traits (e.g. NEO-FFI, emotion regulation, etc.) as well as the current mood. Personality (long-term), action tendency (more medium-term), mood (medium-term) and emotions (short-term) consequently represent different semantic and temporal levels in the emotion model that interact with each other in a realistic fashion.

The special modeling challenge lies in the mapping of non-linear intensity curves (which can differ depending on the emotion or disposition). Aside from the general progression of short-term (emotional) states, the temporal characteristics of changed mid-term to long-term parameters (e.g. mood, action tendency) and/or their reciprocal effect is relevant for predicting emotional behavior. The technical implementation of COMPLEX is shown in Figure 4.

The theoretical basis for the appraisal process currently used in COMPLEX is the OCC model (Ortony et al., 1988). It is based on the assumption that (discrete) emotions are the direct outcome of an individual appraisal process, which appraises an event and/or an action based on three aspects: (1) consequences of the event (for one's own goals), (2) appraisal of the action on the basis of individual standards and (3) certain aspects of objects. These three aspects are further differentiated by the idea of several agents involved in the interaction because the relationship between agents must also be taken into account in the appraisal process. For example, if Person A, who is friends with Person B, fails an exam, which is important to the goal "graduation", the model would generate the emotion *'pity'* for Person A (based on the positive relationship and because an important goal was negatively impacted). For Person C, who has a negative relationship with Person A, the OCC model would, however, predict the emotion *'gloating'*.

To implement these appraisal processes in COMPLEX, it is first of all important to map the individual variables (events/actions, goals, etc.) and their interplay:

Relationships between agents are described by a value ranging from –1 to +1 ($rel_{Ag1,\,Ag2} = [-1, 1]$). The "1 reflects a maximum negative and +1 reflects a maximum positive relationship. The relationship is

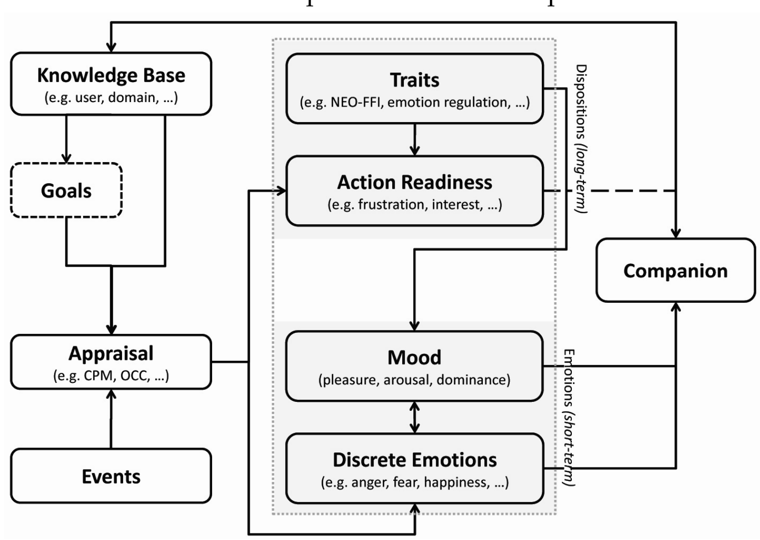

Figure 4. Overview of the COMPLEX model.

directional and therefore does not necessarily apply in the opposite direction. For initialization purposes, if two agents do not know each other, a neutral initial value of 0 is assumed.

The importance of the goal is especially relevant for the individual **goals**. It is represented for each defined goal with a value between 0 (no relevance or non-existent) and 1 (maximum relevance) (rel_{Goal} = [0, 1]).

Currently, COMPLEX only defines the actions of individual agents as **events**. These events are categorized depending on the individual values of each agent with regard to their praiseworthiness ($praiseworthiness_{Action}$ = [–1, 1]). Consequently, this variable defines whether an agent feels "ashamed" for performing an action, or feels "pride" in the other case. The '*praiseworthiness*' variable may range between "1 (low value) and +1 (high value). Internal events (e.g. psychophysiological parameters of the user) are (not yet) taken into account in the current implementation of COMPLEX.

How an event is assessed within COMPLEX also depends on its **consequences** ($consequence_{Goal,\ Group,\ Prospect}$ = [–1, 1]). Consequence is understood as the influence on an individual goal. The *consequence* variable may range between "1 (hindering the goal) and +1 (facilitating the goal). In addition, the consequences of an action are appraised separately for different groups (Group = [Self, Other, Concerned]) and may also occur in the future and not at the time an action is carried out (Prospect = [true, false]). Handing in a paper, for example, may not lead to a feeling of deep satisfaction, but initially only the hope that it is accepted. Satisfaction comes later, once the paper has been accepted.

Consequently, an event or an action may have numerous consequences. The definition of these individual consequences is very difficult to implement in COMPLEX because the variables are domain-specific and must be defined for every application domain. Especially problematic are the variables '*praiseworthiness*' and '*consequence*' because they cannot be defined globally for a domain, but are based on individual values (such as, for example, stable or variable personality traits) and are therefore agent-specific.

The direct output of the appraisal process, which uses the OCC model, are discrete emotions. In terms of time, however, these last milliseconds to seconds are therefore not long enough to model the emotional experience over longer periods of time. COMPLEX tries to solve this problem with the help of additional variables that last longer, for example, moods (medium-term) or personality (long-term). Both the emotional state and the mood are internally mapped as coordinates in the three-dimensional core affect space (valence, aroausal, dominance, VAD). Stable personality traits are taken into account in the current implementation of COMPLEX in a rather rudimentary fashion (variable personality traits are currently not taken into account at all). Depending on the respective personality structure (NEO-FFI), individual starting points are determined in the VAD space for the basic mood (Mehrabian, 1996). These anchor points are firmly located in the VAD space and serve to slowly attract the mood, if no other emotional response is "active" or influences the mood. The speed with which the mood is drifted back to this starting value in the VAD space is described with the help of the formula $I * \cos\left(\sqrt{\frac{d}{m}} * t\right)$ (spring model: I = Starting Intensity, d = Spring Constant, m = Mass, Becker-Asano, 2007). The mood is deflected by the output of the OCC model (emotions). To do so, the discrete emotions of the OCC appraisal are first mapped in the VAP space and averaged. Consequently, the COMPLEX moves both in the discrete and the dimensional space. The result of the averaging is one single point in the VAD space, which serves to manipulate the agent's current mood with regard to valence, arousal and dominance. Taking valence as an example, this is done as follows: $V_{mood\_new} = V_{mood\_old} + \text{neuroticism} * V_{emotion}$. Arousal and dominance are deflected accordingly. A high score on the neuroticism scale, for example, would then lead to a "faster" change of the values within the VAD space, which means that the individual is emotionally less stable.

The modular structure of COMPLEX makes it possible to exchange individual components at any time in new implementations. The OCC model that is currently used for the appraisal process can be replaced by implementation of other theories, such as the component process model (see Figure 2).

It is better to travel alone than with a bad companion. Senegalese

6. Measurements of Emotions, Dispositions and Various Personality Traits

The component process model makes it possible to describe emotions and dispositions in the human-companion interaction in a structural and dynamic manner. This description is initially phenomenological because neither model contains any formal or statistic descriptions of the dynamic relationships between the various components in the process. It is necessary for the operationalization of the individual components to measure and model concrete man-companion interactions (see Table 3). There are, of course, definitions and measuring techniques for all emotional components that lead to scale values and variables. Usually, these operationalizations are not, however, designed for real-time collection. Most suitable are behavior data (video), speech data (audio) and psychobiology data, which can be gathered on a continuous basis during the man-companion interaction. These measurements are not very reactive, i.e. the measurement itself only has little influence on the communicative process. While the measurements themselves are unproblematic, the analysis of the data in real time is a major challenge.

Table 4. Measurements of the emotions and dispositions in the component process model of the emotional behavior based on emotion components.

Measurements Emotions and Dispositions	Subjective Experience	Facial Expressiveness	Psychobiology	Psychomotor Behavior: Gestures, body movements, attention, focus
Moods, core affect	SAM, affect grid, interview	FACS, Ratings, RTAutomatic recognition of facial expressions	Partially through ANS pattern, voice parameters	
Novelty and Valence (N&V)	SAM, affect grid, interview	FACS, EM-FACS, RTAutomatic recognition of facial expressions	P300 (EEG), EDA, EMG, voice parameters	Head movement (OR), defensive reaction, response time, eye-tracking
Discrete Emotions (VAD, DES)	SAM, affect grid, Differential Emotion Scale (DES), interview, semantic	FACS, EM-FACS, ratings, response time	Partially through ANS pattern, voice parameters	Ratings, automatic gesture and body movement detection, localization
Action tendencies	Interview, semantic		Partially through EMG, hemispheric shifts in the EEG power spectrum, voice parameters	Eye-tracking, automatic gesture and body movement detection
Motives	LMI, interview	Rare expressions (contempt)		Eye-tracking
Personality Traits	Personality scales, Interview		Partially through ANS and CNS patterns	

6.1 Measures of subjective experience

The **Self-Affective Manikin** (Bradley and Lang, 1994) can determine the variables valence, arousal and dominance and assess them on a scale from 1 to 9. The rating scale for valence means: "1" is absolutely negative, "5" is neutral and "9" is absolutely positive; for arousal, it means: "1" is absolutely relaxed, "5" is average arousal and "9" is high arousal; and for dominance, "1" means absolute control. The **Differential Emotion Scale** (DES; Izard et al., 1974) consists of 10 emotion categories (interest, joy, sadness, anger, fear, guilt, disgust, surprise, shame, shyness) with three emotion adjectives each.

The **Performance Motivation Inventory** (Schuler and Prochaska, 2000) integrates several dimensions of the performance-oriented personality: Perseverance, dominance, commitment, confidence in success, flexibility, flow, fearlessness, internality, compensatory effort, pride in performance, willingness to learn, preference for difficulties, independence, self-control, status orientation, competition orientation and level of ambition. The analysis may be dimension-specific or as a general value. The results are presented in the form of a profile.

Need for Cognition: The Need for Cognition concept can be individually determined with the German version "Skala zur Erfassung von Engagement und Freude bei Denkaufgaben". The scale is subdivided into three factors: 1. Pleasure from engaging in thinking and solving brain teasers, 2. Positive self-assessment of one's own cognitive abilities and 3. Brooding and conscientiousness (Bless et al., 1994).

The **NEO-FFI** is a proven method to measure five different personality traits: Agreeability, openness, extroversion, neuroticism and conscientiousness. The NEO-FFI consists of 60 questions, 12 for each trait. This model is a data-based, cross-sectional and empirically proven model. The following internal consistencies are provided for the NEO-FFI: Neuroticism = .79, extroversion = .79, openness = .80, agreeability = .75, conscientiousness = .73. The output format is a five-point Likert scale.

The **ERQ** (Abler and Kessler, 2009; based on Gross and John, 2003) makes it possible to scientifically research emotion regulation processes. Preferences for two frequently used strategies can be identified: suppression and reappraisal. To determine these two parameters, 10 items each are provided, which can be rated from 1 ("do not agree at all") to 7 ("agree completely"). The German version reaches an internal consistency of r = .74 for "suppression" and r = .76 for "reappraisal".

6.2 Measures of facial expressiveness

Human codings: The most commonly used system is the Facial Action Coding System, an anatomy-based method to describe visually distinguishable facial movements. The FACS does not interpret the facial expression, but detects inseparable facial movements, the so-called 44 facial action units (AUs). Individual muscles create the basis for AUs, but some AUs are produced by groups of muscles. Some muscles can even produce several different AUs. With the help of the 44 AUs, it is possible to describe all emotional (EmFACS) and non-emotional movements (such as communication signals) of the face. The analysis is based on a video recording. Furthermore, the intensity and temporal resolution can be captured (onset, apex, offset).

The Active Appearance Model (Cohn, 2010) and CERT (Bartlett et al., 2008) are able to automatically recognize the AUs. Such automated video processing systems for the recognition of facial movements are based on the gathering and classification of features (Wimmer and Radig, 2007). The efficiency of these methods is significantly increased by taking into account the modeling of the facial geography and thus the separation of dynamic (contours, coloring and contrasts) and static features (Skelley et al., 2006).

6.3 Measures of psychobiological activity

The psychobiological emotion recognition can be subdivided into traditional psychological research (Stemmler and Wacker, 2010; Kreibig et al., 2007) and the area of affective computing (Picard et al., 2001). It has been noticed, however, that the two areas are converging (Kolodyazhniy et al., 2011; Walter et al., 2010; Hrabal et al., 2012). In the basic research area and the affective computing area, discrete (fear, anger, joy, etc.) and dimensional (valence, arousal, dominance) models are used. In both research areas, however, the dimensional approach has become more prevalent. Both areas use similar parameters: blood volume pulse (BVP), skin conductance level (SCL), respiration (RSP) and electromyography (EMG). Von Kolodyazhniy and coworkers (2011) added additional parameters. Especially reliable are the correlations between the dimensions valence and corrugator and/or Zygomaticus EMG (Tan et al., 2012), as well as between arousal and SCL or BVP, respectively. In this process, predominantly action potential and frequency analyses are determined. Frequency analyses are generally suitable for real-time applications.

The main difference between basic research and affective computing is the following: In the basic research domain, averages and standard deviations are formed for signals, and signal increases or decreases calculated with these. To do so, statistical methods such as *t*-tests and variance analysis are used and effect strengths calculated (Kreibig et al., 2007, 2010; Stemmler and Wacker, 2010; Schupp et al., 2004; Bradley, 2009). In the affective computing area, generally spoken, the raw signals are initially subjected to a (1) **pre-processing** and then a (2) **feature extraction** ($f_{\#}$). The number of extracted features ranges from 13 (Haag et al., 2004) to 110 (Kim and André, 2008).

There is no consensus yet about which feature extraction is better. As an example, the extraction of Gu et al. (2008) will be mapped. Gu et al. used the following formulas for corrugator and Zygomaticus EMG, BVP, SCL, temperature (TMP) and electrocardiogram parameters (ECG) for each of the six parameters and extracted 36 features:

- The mean of $x(n)$
- The standard derivation of $x(n)$
- The mean of the absolute values of the first differences of $x(n)$
- The mean of the absolute values of the first differences of normalized $x(n)$
- The mean of the absolute values of the second differences of $x(n)$
- The mean of the absolute values of the second differences of the normalized $x(n)$

What follows is an automatic (3) **feature selection**. Gu et al. (under review) were able to show with the Sequential Floating Forward Search (SFFS) algorithm that, with regard to valence, the accuracy and robustness reaches their highest levels at 10 features and starts to decrease at 20 features, but arousal already reaches its highest level at 5 features and starts to decrease at 22 features. The problem with the feature selection is that these features are individual-specific and trans-situational dependent. Nevertheless, Kolodyazhniy et al. (2011) selected features that are inter-individually and trans-situationally robust: SCL, Corrugator-EMG, Zygomathicus-EMG, pCO_2 (end-tidal carbon dioxide partial pressure) and PEP (pre-ejection period).

The last step is the (4) **classification** (LDP, SVM, MLP, etc.) or hybrid classification, respectively. Overall, however, based on current findings, it can be said that psychobiological signals have the advantage that they can be obtainable in a permanent fashion and regardless of the location. It is absolutely necessary, however, that sensors are "comfortable". The psychobiological gathering of data tends to require individual-specific processing (Walter et al., 2013, Böck et al., 2012). A trans-situational, robust feature selection still currently presents a problem (Walter et al., 2013).

6.4 Measurements of prosody and paralinguistic parameters

Verbal communication can be analyzed in a hierarchically ordered manner. After processing the acoustic information with a signal-theoretical approach, prosodic and linguistic language information is processed separately. The prosodic information contains references about the emotional and motivational content of the act of communication. These correspond with or differ from the linguistic emotion- and cognition-related information from the semantics. On the highest level, the prosody and semantics of the spoken information can be interpreted as possible intentions. Within the context of this framework, reference is made only to emotion-related information.

The analysis of emotional language produces a multitude of prosodic characteristics, which differ significantly from modal and/or unemotional language. Among others, features such as fundamental frequency progressions, sound pitch progressions, energy progressions, speech pause frequency and pattern, the stretching of words and syllables, frequency changes and voice qualities were identified (Scherer, 2003a; Yanushevskaya et al., 2007). Researchers working in the area of automatic emotion recognition in speech use these features (Borst et al., 2004). In addition, there are many para-linguistic events that transmit emotionally colored information, such as laughing (Scherer et al., 2011). As already mentioned above, semantic and speech-content phenomena are also researched in addition to the prosodic and para-linguistic information hidden in speech (Schuller et al., 2003; Scherer et al., 2012b).

6.5 Semantics

Content-analytical processes, for some time now, have been based on key words or phrases. More complex processes carry out syntactic and semantic analyses. In certain areas of application with a limited speech scope, the results are robust and can be applied in practice. The text-based detection of emotional information is a proven method: Harvard Dictionaries, LIWC (Pennebaker and Francis, 1996) and LEAS (Kessler et al., 2010) and is easy to use, once words have been identified. Since natural speech rarely follows grammar rules, computer-linguistic methods often face significant problems in the identification of meanings. The automated content analysis can be rendered less ambiguous with decision-theoretical methods.

So far, however, there is an insufficient number of language databases with natural emotionality in the language. First analyses, however, show promising emotion classification methods (Gnjatović and Rösner, 2010). Emotion-related annotated data is classified through a dynamic recognition process based on Hiden Markov models, neuronal networks, and kernel-based processes as well as support vector machines (SVM) and kernel logistic regression (KLR). The language-independent classification of basic emotions lies at approximately 75%, and the language-dependent classification at 93%, which is very precise, with regard to artificially created language material with emotional prosody (Wagner, 2005), if the language signals come from a defined interaction and are recorded without any acoustic interference.

6.6 Psychomotor functions: Gestures, body movements and attention focus

The automatic recognition of gestures generally takes place in three steps: (1) the object detection, (2) the chronologically recursive filtering (tracking) and (3) the classification and/or verification (Bar-Shalom and Li, 1995). (1) In the detection step, the measured data is individualized, which means that object hypotheses with individually measured data are extracted from the measured data, which then must be allocated to the known objects. (2) The chronological filtering then, in the second step, uses a multi-instance filter approach, in which every object is individually assigned a dynamic filter. For chronological filtering processes, approximations of the generally recursive Bayes estimation are used (Bar-Shalom and Li, 1995). (3) In the area of gesture recognition, numerous classification methods have been developed ((Morguest, 2000), (Corradini and Gross, 2000) and (Barth and Herpers, 2005)). Once the static and dynamic gestures have been recognized, they can be classified. For static gestures, model-based processes (Stenger et al., 2001), Active Shape Models (Wimmer and Radig, 2007), or feature-based processes in connection with a classification algorithm (e.g. Bayesian Networks (Ong and Ranganath, 2003) or artificial neuronal network, 2001) can be used. In the dynamic methods, the focus is on the analysis of a chronological sequence of individual images to determine the movement trajectory. To do so, image and pattern recognition methods for time-dependent features must be used in order to analyze the information from the video sequence. The typical sequential process flow for gesture classification includes segmentation, feature recognition and feature extraction.

Good company makes short miles. Dutch.

7. Multi-modal Assessment and Multi-modal Fusion for Emotion and Disposition Recognition

Recognizing the users' emotional and dispositional states can be achieved by analyzing different modalities, e.g. analyzing facial expressions, body postures, and gestures, or detecting and interpreting paralinguistic information hidden in speech (see section on measurements above). In addition to these types of external signals, psychobiological channels can provide information about the user's current emotional state (honest signals sensu Pentland and Pentland, 2008). Although emotion recognition is often performed on single modalities, particularly in benchmark studies on acted emotional data, for the recognition of more naturalistic emotions, multi-modal events or states need to be considered, and principles of multi-modal pattern recognition become increasingly popular (Caridakis et al., 2007; Walter et al., 2011).

Basically, any multi-modal classification problem can be treated as a uni-modal one, just by extracting relevant data or feature vectors from each modality and concatenating them in a single vector that is then applied as input vector to a single monolithic classifier. This fusion scheme is called *data fusion*, *early fusion* or *low-level fusion*. The opposite of data fusion is *decision fusion*, *late fusion* or *high-level fusion*. This means that information of different modalities is processed separately until the classifier decisions were computed. After that an aggregation rule is applied combining the entire bunch of decisions into a final overall decision. All these different notions are reflecting the processing level (data/decision, early/late, or low/high) where information fusion takes place. In addition to these two principles, *feature (level) fusion* or *intermediate (level) fusion* or *mid (level) fusion* is a common fusion scheme. This notion is used to express the fact that information sources are fused after computing some type of higher-level discriminative features, e.g. action-unit intensities, statistics of spoken words, speech content (Schwenker et al., 2006).

Besides the spatial fusion types of different modalities, in multi-modal data streams the integration of temporal information is required (Dietrich et al., 2003). In human-computer interaction scenarios of typical events in the environment, the user's states or actions cannot be detected or classified on the basis of single video frames or short-time speech analysis windows (Glodek et al., 2011b). Usually such events or states are represented through multi-variate time series and thus fusion in these applications almost always means both spatial and temporal information fusion. The simplest temporal fusion scheme is chunking by filtering, averaging, or static decision fusion such as (weighted) voting. Chunking assumes that entries of the series are independent and can be computed separately; this assumption might be true for the classification of a global emotional state, but in case of data, such as actions or user dispositions, the sequential nature has to be explicitly modeled through recurrent neural networks or hidden Markov models (see also Chapter 4 in this book; Bishop, 2006; Glodek et al., 2011a).

REFERENCES

Abler, B. and H. Kessler. 2009. Emotion Regulation Questionnaire—Eine Deutsche Version des ERQ von Gross and John. *Diagnostica*, **55(3):** 144–152.

Antonovsky, A. 1987. Unraveling the Mystery of Health—How People Manage Stress and Stay Well. Jossey-Bass Publishers, San Francisco.

Asendorpf, D. 2011. Netz auf Rädern. DIE ZEIT, March 10th, 40 p.

Aubrey, B. 2010. Managing Your Aspirations: Developing Personal Enterprise in the Global Workplace. McGraw-Hill Education, Singapore.

Bandura, A. 2001. Social cognitive theory: An agentic perspective. *Ann. Rev. Psychol.*, **52:**1–26.

Bar-Shalom, Y. and X.R. Li. 1995. Multitarget-Multisensor Tracking—Principles and Techniques. YBS, Urbana, IL.

Barth, A. and R. Herpers. 2005. Robust head detection and tracking in cluttered workshop environment using GMM, pp. 442–450. In DAGM symposium, LNCS 3663. Springer, New York, Berlin.

Bartlett, M. G. Littlewort, E. Vural, K. Lee, M. Cetin, A. Ercil and M. Movellan. 2008. Data mining spontaneous facial behavior with automatic expression coding. Lecture Notes in Computer Science 5042, Springer, Berlin, pp. 1–21.

Baum, S.R. and M.D. Pell. 1999. The neural bases of prosody: Insights from lesion studies and neuroimaging. *Aphasiology*, **13:**581–608.

Becker-Asano, C. 2008. WASABI: Affect Simulation for Agents with Believable Interactivity. Dissertation der Universität Bielefeld.

Bierhoff, H.-W. 2011. Attribution: In Lexikon der Psychologie. http://www.wissenschaft-online.de/abo/lexikon/psycho/1584.

Bishop, C. 2006. Pattern Recognition and Machine Learning. Springer, New York, Berlin.

Biundo, S. and A. Wendemuth. 2012. A Companion Technology for Cognitive Technical Systems. Proceedings of the EUCogII-SSPnet-COST2102 International Conference, Dresden. In: A. Vinciarelli, R. Hoffman, V.C. Müller (eds.), *Behavioural Cognitive Systems*, LNCS/LNAI series. Springer, Berlin.

Biundo, S. and A. Wendemuth. 2010. Von kognitiven technischen Systemen zu Companion-Systemen. *Künstliche Intelligenz*, **24(4):**335–339.

Böck, R., K. Limbrecht, S. Walter, S. Glüge, D. Hrabal, A. Wendemuth and H.C. Traue. 2012. Intraindividual and interindividual multimodal emotion analyses in human-machine-interaction. Proceedings of 2012 IEEE Conference on Cognitive Methods in Situation Awareness and Decision Support, pp. 59–64.

Bless, H., M. Wänke, G. Bohner and R.F. Fellhauer. 1994. Need for Cognition: Eine Skala zur Erfassung von Engagement und Freude bei Denkaufgaben. *Zeitschrift für Sozialpsychologie*, **25:**147–154.

Borst, M., G. Langner and G. Palm. 2004. A biologically motivated neural network for phase extraction from complex sounds. *Biological Cybernetics*, **90:**98–104.

Bradley, M.M. 2009. Natural selective attention: Orienting and emotion. *Psychophysiology*. **46(1):**1–11.

Bradley, M.M. and P.J. Lang. 1994. Measuring emotion: The self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry*, **25:**49–59.

Bradley, M.M. and P.J. Lang. 2008. The International Affective Picture System (IAPS) in the study of emotion and attention, pp. 29–46. In: J.A. Coan and J.J. Allen (eds.), Handbook of Emotion Elicitation and Assessment. Oxford University Press, Oxford.

Bryson, J.J. 2010. Robots should be slaves, pp. 63–74. In: Y. Wilks (ed.), Close Engagements with Artificial Companions: Key Social, Psychological, Ethical and Design Issues. John Benjamins, Amsterdam.

Caridakis, G., G. Castellano, L. Kessous, A. Raouzaiou, L. Malatesta and S. Asteriadis. 2007. Multimodal emotion recognition from expressive faces, body gestures and speech, pp. 375–388. In: C. Boukis, L. Pnevmatikakis and L. Polymenakos (eds.), Artificial Intelligence and Innovations 2007: From Theory to Applications. Springer, Boston, Berlin.

Cohn, J.F. 2010. Advances in behavioral science using automated facial image analysis and synthesis. IEEE Signal Processing Magazine, **128(6):**128–133.

Colombetti, G. 2009. From affect programs to dynamical discrete emotions. *Philosophical Psychology*, **22(4):**407–445.

Corradini, A. 2001. Real-time gesture recognition by means of hybrid recognizers, pp. 34–46. Proceedings of the International Gesture Workshop on Gesture and Sign Languages in Human-Computer Interaction, USA.

Corradini, A. and H. Gross. 2000. Camera-based gesture recognition for robot control, pp. 133–138. In Proceedings of IJCNN'2000.

Costa, P.T. and R.R. McCrae. 2002. Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI) 2002. Psychological Assessment Resources, Odissa.

Cowie, R. and R.R. Cornelius. 2003. Describing the emotional states that are expressed in speech. *Speech Communication*, **40:**5–32.

Deighton, R.M. and H.C. Traue. 2006. Emotionale Ambivalenz, Körperbeschwerden, Depressivität und soziale Interaktion: Untersuchungen zur deutschen Version des Ambivalence over Emotional Expressiveness Questionnaire (AEQ-G18. *Zeitschrift für Gesundheitspsychologie*, **14:**158–170.

Dietrich, C., G. Palm and F. Schwenker. 2003. Decision templates for the classification of bioacoustic time series. *Information Fusion*, **4:**101–109.

Ekman, P. and W. Friesen. 1978. *Facial Action Coding System: Investigator's Guide*. Palo Alto: Consulting Psychologists Press.

Frick, R.W. 1985. Communicating emotion: The role of prosodic features. *Psychol. Bull.*, **97:**412–429.

Friederici, A.D. and K. Alter. 2004. Lateralization of auditory language functions: A dynamic dual pathway model. *Brain Lang*, **89:**267–276.

Frijda, N.H. 1988. The laws of emotion. *Am. Psychol.,* **43:**349–358.

Frijda, N.H. 2007. The Laws of Emotion. Erlbaum, Mahwah New Jersey and London.

Frijda, N.H. and M. Zeelenberg. 2001. Appraisal: What is the dependent? pp. 141–155. In: K.R. Scherer, A. Schorr and T. Johnstone (eds.), Appraisal Processes in Emotion: Theory, Methods, Research. Oxford University Press, New York.

Garcia-Prieto, P. and K.R. Scherer. 2006. Connecting social identity theory and cognitive appraisal theory of emotions, pp. 189–208. In: R. Brown and D. Capozza (eds.), Social Identities: Motivational, Emotional, Cultural Influences. Psychology Press, New York.

Glodek, M., L. Bigalke, M. Schels and F. Schwenker. 2011a. Incorporating uncertainty in a layered HMM architecture for human activity recognition, pp. 33–34. Proceedings of the 2011 Joint ACM Workshop on Human Gesture and Behavior Understanding. 2011. ACM.

Glodek, M., S. Tschechne, G. Layher, M. Schels, T. Brosch, S. Scherer, M. Kächele, M. Schmidt, H. Neumann, G. Palm and F. Schwenker. 2011b. Multiple classifier systems for the classification of audio-visual emotional states, pp. 359–368. In: S.D'Mello et al. eds.)], ACII 2011, Part II, LNCS 6975, Springer, Heidelberg.

Gnjatović, M. and D. Rösner. 2010. Inducing genuine emotions in simulated speech-based human-machine interaction: The NIMITEK Corpus. *IEEE Transactions on Affective Computing,* **1:**132–144.

Gross, J.J. and O.P. John. 2003. Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *J. Pers. Soc. Psychol.,* **85:**348–362.

Gu, Y., S.L. Tan, K.J. Wong, M.H.R. Ho and L. Qu. 2008. Using GA-based feature selection for emotion recognition from physiological signals. International AH '10 Proceedings of the 1st Augmented Human International Conference. New York, NY: ACM, pp. 1–4.

Guralnik, D.B. 1970. *Webster's New World Dictionary*. Foster, Toronto/Canada.

Haag, A., S. Goronzy, P. Schaich and J. Williams. 2004. Emotion recognition using biosensors: First step towards an automatic system. Affective Dialogue Systems, Tutorial and Research Workshop, Kloster Irsee, Germany, pp. 36–48.

Hrabal, D., S. Rukavina, K. Limbrecht, S. Walter, V. Hrabal, S. Gruss and H.C. Traue. 2012. Emotion Identification and Modelling on the Basis of Paired Physiological Data Features for Companion Systems. Proceedings of the MATHMOD 2012—7th Vienna International Conference on Mathematical Modelling. Vienna, Austria.

Izard, C.E., F.E. Dougherty, B.M. Bloxom and N.E. Kotsch. 1974. The Differential Emotions Scale: A Method of Measuring the Meaning of Subjective Experience of Discrete Emotions. Nashville: Vanderbilt University, Department of Psychology.

James, T.W., G.K. Humphrey, G.S. Gati, S.J.S. Ravi, R.S. Menon and M.A. Goodale. 2000. The effects of visual object priming on brain activation before and after recognition. *Current Biol.,* **10:**1017–1024.

Keller, J., S. Ringelhan and F. Blomann, 2011. Does skills-demands compatibility result in intrinsic motivation? Experimental test of a basic notion proposed in the theory of flow-experiences. *Journal of Positive Psychology*, **6(5):**408–417.

Kernberg, O.F. 1992. Objektbeziehungen und Praxis der Psychoanalyse. Klett-Cotta Stuttgart.

Kessler, H., A. Festini, H.C. Traue, S. Filipic, M. Weber and H. Hoffmann. 2008. SIMPLEX—Simulation of Personal Emotion Experience, Affective Computing. pp. 255–271. In: Jimmy Or (ed.), Affective Computing: Emotion Modelling, Synthesis and Recognition, I-Tech Education and Publishing.

Kessler, H., H.C. Traue, M. Hopfensitz, C. Subic-Wrana and H. Hoffmann. 2010. Level of Emotional Awareness Scale-Computer (LEAS-C): Deutschsprachige digitale Version. *Psychotherapeut, Bnd 55, Heft,* **4:**329–331.

Kim, J. and E. André. 2008. Emotion recognition based on physiological changes in music listening., *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **30:**2067–2083.

Kohut, H. 1987. The Kohut Seminars on Self Psychology and Psychotherapy with Adolescents and Young Adults. W.W. Norton and Company, New York.

Kolodyazhniy, V., S. Kreibig, J.J. Gross, W. Roth and F.H. Wilhem. 2011. An affective computing approach to physiological emotion specificity: Toward subject-independent and stimulus-independent classification of film-induced emotions. *Psychophysiology*, **48:**908–922.

Krapp, A. 2002. Structural and dynamic aspects of interest development: theoretical considerations from an ontogenetic perspective. Learning and Instruction. **12:**383–409.

Kreibig, S.D., F.H. Wilhelm, W.T. Roth and J.J. Gross. 2007. Cardiovascular, electrodermal, and respiratory response patterns to fear and sadness-inducing films. *Psychophysiology*, **44:**787–806.

Kreibig, S.D., T. Brosch and G. Schaefer, 2010. Psychophysiological response patterning in emotion: Implications for affective computing, pp. 105–130. In: K.R. Scherer, T. Baenziger and E. Roesch (eds.), A Blueprint for an Affectively Competent Agent: Cross-Fertilization Between Emotion Psychology, Affective Neuroscience, and Affective Computing. Oxford University Press, Oxford.

Kuhl, J. 2001. Motivation und Persönlichkeit. Interaktionen psychischer Systeme. Hogrefe, Göttingen.

Leventhal, H. and K. Scherer. 1987. The relationship of emotion to cognition: a functional approach to semantic controversy. *Cognition and Emotion*, **(1):**3–28.

Loewenstein, G. and J.S. Lerner, 2003. The role of affect in decision making, pp. 619–642. In: R.J. Davidson, K.R. Scherer and H.H. Goldsmith (eds.), Handbook of Affective Sciences. Oxford University Press, Oxford.

Maruta, T., R.C. Colligan, M. Malinchoc and K.P. Offord. 2002. Optimism-pessimism assessed in the 1960s and self-reported health status 30 years later. *Mayo. Clin. Proc.*, **77:**748–753.

Maslow, A.H. 1943. A theory of human motivation. *Psychol. Rev.*, **50:**370–396.

Mehrabian, A. 1996. Analysis of the big-five personality factors in terms of the PAD temperament model. *Austral. J. Psychol.*, **48:**86–92.

Miller, N.E. 1959. Liberalization of basic S-R concepts: Extension to conflict behavior, motivation, and social learning. In: S. Koch (ed.), Psychology: A Study of Science (Vol. 2. McGraw-Hill, New York).

Morguet, P. 2000. Stochastische Modellierung von Bildsequenzen zur Segmentierung und Erkennung dynamischer Gesten. Ph.D. thesis, TU München, München, Germany.

Nass, C. and S. Brave. 2005. Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship. MIT Press, Cambridge, MA.

Ong, S. and S. Ranganath, 2003. Classification of gesture with layered Meanings. Proceedings of International Gesture Workshop Singapore, pp. 239–246.

Ortony, A., G.L. Clore and A. Collins. 1988. The Cognitive Structure of Emotions. Cambrige: Cambrige University Press.

Pennebaker, J.W. and M.E. Francis. 1996. Cognitive, emotional, and language processes in disclosure. *Cogn. Emot.*, **10:**601–626.

Pentland, A. and S. Pentland. 2008. Honest Signals: How They Shape Our World. Bradford Books. Cambrige: MIT Press.

Picard, R. 2003. Affective computing: Challenges. *International Journal of Human-Computer Studies*, **59:**55–64.

Picard, R.W., E. Vyzas and J. Healey. 2001. Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **23:**1175–1191.

Puchalska-Wasyl, M. 2007. Types and functions of inner dialogues. *Psychology of Language and Communication*, **11:**43–62.

Scherer, S., M. Glodek, M. Schels, M. Schmidt, G. Layher, F. Schwenker, et al. 2012a. A generic framework for the inference of user states in human computer interaction: How patterns of low level communicational cues support complex affective states. *Journal on Multimodal User Interfaces*, **6:**117–141.

Scherer, S., M. Glodek, F. Schwenker, N. Campbell and G. Palm. 2012b. Spotting laughter in naturalistic multiparty conversations: A comparison of automatic online and offline approaches using audiovisual data. ACM Transactions on Interactive Intelligent Systems: Special Issue on Affective Interaction in Natural Environments, **2:**111–144.

Scherer, S., E. Trentin, F. Schwenker and G. Palm. 2009. Approaching emotion in human computer interaction. In: International Workshop on Spoken Dialogue Systems. (IWSDS'09), pp. 156–168.

Scherer, K.R. 2001. Appraisal considered as a process of multilevel sequential checking, pp. 92–120. In: K.R. Scherer, A. Schorr and T. Johnstone (eds.), Appraisal Processes in Emotion: Theory, Methods, Research. Oxford University Press, New York.

Scherer, K.R. 2003a. Vocal communication of emotion: A review of research paradigms. *Speech Communication*, **40:**227–256.

Scherer, K.R. 2003b. Introduction: Cognitive components of emotion, pp. 563–571. In: R.J. Davidson, K.R. Scherer and H.H. Goldsmith (eds.), Handbook of Affective Sciences. Oxford University Press, Oxford.

Schuler, H. and M. Prochaska. 2000. Entwicklung und Konstruktvalidierung eines berufsbezogenen Leistungsmotivationstests. *Diagnostica*, **46:**61–72.

Schuller, B., G. Rigoll and M. Lang, 2003. Hidden Markov model-based speech emotion recognition. *Proceedings IEEE International Conference on Multimedia and Expo*, pp. 401–404.

Scherer, S., M. Schels and G. Palm. 2011. How low level observations can help to reveal the user's state in hci. In S. D'Mello, A. Graesser, B. Schuller, and J.-C. M. (eds.) Proceedings of the 4th International conference on Affective Computing and Intelligent Interaction (ACII'11), volume 2, pp. 81-90. Berlin: Springer .

Schupp, H.T., B.N. Cuthbert, M.M. Bradley, C.H. Hillman, A.O. Hamm and P.J. Lang, 2004. Brain processes in emotional perception: Motivated attention. *Cogn. Emot.*, 18:593–611.

Schwenker, F., C. Dietrich, C. Thiel and G. Palm. 2006. Learning of decision fusion mappings for pattern recognition. *International Journal on Artificial Intelligence and Machine Learning (AIML)*, 6:17–21.

Skelley, J., R. Fischer, A. Sarma and B. Heisele. 2006. Recognizing expressions in a new database containing played and natural expressions. *Proceedings of 18th International Conference on Pattern Recognition (ICPR)*, **1:**1220–1225.

Sloman, A. 2010. Requirements for artificial companions: It's harder than you think, pp. 179–200. In: Y. Wilks (ed.), Close Engagements with Artificial Companions. J. Benjamins Publishing Company, Amsterdam.

Stemmler, G. and J. Wacker. 2010. Personality, emotion, and individual differences in physiological responses. *Biol. Psychol.*, **84:**541–551.

Stenger, B., P. Mendonca and R. Cipolla. 2001. Model-based 3D tracking of an articulated hand. *IEEE Conference on Computer Vision and Pattern Recognition*, **2:**310–315.

Tan, J.-W., S. Walter, A. Scheck, D. Hrabal, H. Hoffmann, H. Kessler and H.C. Traue. 2012. Repeatability of facial electromyography (EMG) activity over corrugator supercilii and zygomaticus major on differentiating various emotions. *Journal of Ambient Intelligence and Humanized Computing*, **3:**3–10.

Traue, H.C. and R.M. Deighton. 2007. Emotional inhibition, pp. 908–913. In: George Fink (Editor-in-Chief), Encyclopedia of Stress, Second Edition, Volume 1. Academic Press, Oxford.

Traue, H.C. and H. Kessler. 2003. Psychologische Emotionskonzepte, pp. 20–33. In: A. Stephan and H. Walter (eds.), Natur und Theorie der Emotion. mentis Verlag, Paderborn.

Traue, H.C., H. Kessler and A.B. Horn. 2005. Emotion, Emotionsregulation und Gesundheit, pp. 149–171. In: R. Schwarzer (ed.), Gesundheitspsychologie. Enzyklopädie der Psychologie. Hogrefe, Göttingen.

Turtle, S. 2010. In good company, pp. 3–10. In: Y. Wilks (ed.), *Close Engagements with Artificial Companions*. J. Benjamins Publishing Company, Amsterdam.

Wagner, J. 2005. Vom physiologischen Signal zur Emotion: Implementierung und Vergleich ausgewählter Methoden zur Merkmalsextraktion und Klassifikation. M.S. thesis, Universität Augsburg, Augsburg, Germany.

Wagner, J., J. Kim and E. André. 2008. From physiological signals to emotions: implementing and comparing selected methods for feature extraction and classification. *Proceedings of IEEE International Conference on Multimedia and Expo ICME,* pp. 940–943.

Walter, S., D. Hrabal, A. Scheck, H. Kessler, G. Bertrain, F. Nodtdurft and H.C. Traue, 2010. Individual emotional profiles in Wizard-of-Oz-Experiments. In: F. Makedon, I. Maglogiannis and S. Kapidakis (eds.), *Proceedings of the 3rd International Conference on Pervasive Technologies Related to Assistive Environments*. **S:**58–63.

Walter, S., S. Scherer, M. Schels, M. Glodek, D. Hrabal, M. Schmidt, R. Böck, K. Limbrecht, H.C. Traue and F. Schwenker. 2011. Multimodal Emotion Classification in Naturalistic User Behavior. In J.A. Jacko (ed.). Human Computer Interaction, Part III, HCI 2011, LNCS 6763, Springer-Verlag Berlin, pp. 603–611.

Walter, S.; Jonghwa Kim; Hrabal, D.; Crawcour, S.C.; Kessler, H.; Traue, H.C., "Transsituational Individual-Specific Biopsychological Classification of Emotions," Systems, Man, and Cybernetics: Systems, IEEE Transactions on, vol. 43, no. 4, pp. 988–995, July 2013. doi: 10.1109/TSMCA.2012.2216869

Weiner, B. 1986. An Attributional Theory of Motivation and Emotion. Springer, New York, Berlin.

Wilks, Y. 2010. Introducing artificial companions, pp. 11–22. In: Y. Wilks (ed.), Close Engagements with Artificial Companions. J. Benjamins Publishing Company, Amsterdam.

Wimmer, M. and B. Radig. 2007. Automatically Learning the Objective Function for Model Fitting. Proceedings of the Meeting in Image Recognition and Understanding (MIRU) in Hiroshima, Japan.

Witherington, D.C. and J.A. Crichton. 2007. Framework for understanding emotions and their development: Functionalist and dynamic systems approaches. *Emotion*, **7:**628–637.

Woolf, B.P. 2010. Social and caring tutors: ITS 2010 Keynote Address, pp. 5–14. In: V. Aleven, J. Kay and J. Mostow (eds.), *Intelligent Tutoring Systems, Part I, LNCS 6094*. Springer, Berlin, Heidelberg.

Yanushevskaya, I., M. Tooher, C. Gobl and A. Ní Chasaide. 2007. Time- and amplitude-based voice source correlates of emotional portrayals, pp. 159–170. In: A. Paiva, R. Prada and R.W. Picard (eds.), Affective Computing and Intelligent Interaction: Proceedings of the ACII 2007 (Vol. 4738). Springer-Verlag, Lisbon, Portugal.

Zajonc, R.B. 1980. Thinking and feeling: Preferences need no interferences. *Am Psychol*, **35:**151–175.