

Evaluating Emotion Expressing Robots in Affective Space

Kolja Kühnlenz, Stefan Sosnowski and Martin Buss
*Institute of Automatic Control Engineering, Technische Universität München
 Germany*

1. Introduction

Research on human emotions has been an area of increased interest in the field of human-robot interaction in the last decade. Subfields reach from usability studies over emotionally enriched communication to even social integration in human-robot groups. Prominent aims are the investigation of the impact of emotional responses, perception of emotions, and emotional decision making on the efficiency and robustness of the interaction process. Intuitive communication and easy familiarization are other factors of major interest.

In order to facilitate emotionally enriched communication, means of expressing “emotional states” of a robot are necessary, i.e. expressive features, which can be used to induce emotions in the human or simply to provide additional cues on the progression of the communication or interaction process. A common approach is the integration of facial expression elements in the robot artefact as very elaborated frameworks on human facial expressions exist, which can be utilized, e.g. (Blow et al., 2006; Breazeal, 2002a; Grammer & Oberzaucher, 2006; Hara & Kobayashi, 1996; Sosnowski et al., 2006a; Zecca et al., 2004).

The design and control of such expressive elements have a significant impact on how the represented emotional state of the robot is perceived by the human counterpart. Particularly, the controlled posture is an important aspect and a well investigated issue in human nonverbal communication considering facial expressions. Common frameworks are works using the Facial Action Coding System (FACS) (Ekman & Friesen, 1977) and variants establishing the link between muscular activations and facial expressions, i.e. the quantitative contribution of muscular group poses to perceived emotions, e.g. (Grammer & Oberzaucher, 2006). Such a design approach is dimensional (continuous) in nature as a continuous representation of the emotional state space composed of the dimensions valence/pleasure, arousal, and dominance/stance is used and the contribution of muscular group poses to these components is provided.

The choice of concept for the evaluation of displayed facial expressions is an issue of equal importance. A comprehensive evaluation is essential as the actuating elements of the robot (motors, joints, transmission elements, etc.) differ significantly from those of the human. Thus, although elaborated frameworks as e.g. FACS are used in the design process a significant deviation of the intended and perceived expression can be expected. Common evaluation procedures use a categorical approach where test participants may choose best fits from a set.

Source: Human-Robot Interaction, Book edited by Nilanjan Sarkar,
 ISBN 978-3-902613-13-4, pp.522, September 2007, Itech Education and Publishing, Vienna, Austria



Figure 1. Robot head EDDIE with actuated features for displaying facial and non-facial expressions (Sosnowski et al., 2006a)

Design and evaluation are, thus, commonly conducted using different models. In consequence evaluations are of low test-theoretical validity as a psychological test not only examines the subject but also the theory in which the subject is embedded. Another shortcoming is the lack of feedback for design improvements, e.g. guidelines for posture adjustments of the expressive features.

This chapter discusses the use of dimensional approaches for evaluation of emotion expressing robots. In the first part a theoretical framework is established and guidelines for evaluation are derived. In the second part these guidelines are exemplarily implemented for illustration purposes and two user studies are presented evaluating the robot head EDDIE.

An additional contribution of the framework discussed in the second part is the use of dimensional evaluation approaches as a generic tool for integrating expressive elements of arbitrary design (e.g. animal-like) for which no common framework as e.g. FACS exists. This is illustrated by a third pilot user study evaluating the impact of EDDIE's ears and crown on the dimensions of the emotion model.

The chapter is organized as follows: Section 2 introduces and discusses common dimensional measures exemplified in the field of emotional expressions; Section 3 presents the application of a semantic differential approach to the evaluation of a facial expression robot head; conclusions are given in Section 4.

2. Dimensional Evaluation Approaches

2.1 Introduction of Quality Measures for Tests

By definition, a test is a specific psychological experiment. The goal of this experiment is to obtain comparative judgments about different subjects and their attitudes, impressions, or psychological and physiological variables (Ertl, 1965). The variables measured in a subject can be divided into two groups: latent and manifest variables. Manifest variables are easily observable like the height or the weight of a person. Latent variables like attitudes, feelings or personal traits are not directly observable and, thus, have to be derived from the answering behavior in a test. The idea of using a test in order to obtain information of a latent variable is depicted in Figure 2.

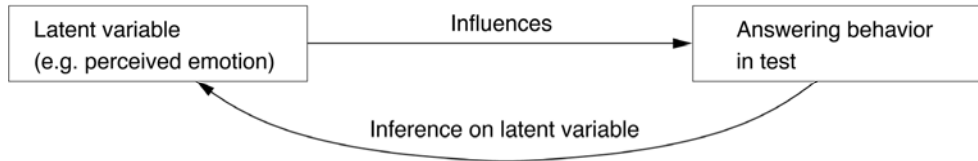


Figure 2. Correlation of answering behavior in a test and latent variables

Thereby, it is assumed that the latent variable influences the answering behavior in a test. After having obtained the answers, the latent variable is deduced from the observed answers of the test (mostly a questionnaire or an interview). This deduction is the most difficult part of the construction of a test since the correlation of latent variables and answering behavior in the test cannot be formulated easily. For that reason, it is very important that the deduction of the latent variable shall be embedded in a profound theory of how true feelings and attitudes of people can be obtained by using tests. In the following paragraph the most common and mostly used theory of measuring feelings and attitudes is described.

2.2 The Semantic Differential

Osgood et al. tried to connect scaled measurement of attitudes to the connotative meaning of words (Osgood et al., 1957). In their classical experiment they discovered that the whole semantic space of words can be described by just three dimensions. These dimensions are evaluation (e.g. good vs. bad), potency (e.g. strong vs. weak), and activity (e.g. aroused vs. calm). The measurement technique they used is called semantic differential (approx. 20-30 bipolar adjectives on a seven-point Likert-scale). By using this method, a person's attitude can be plotted in a semantic space and different subjects' attitudes towards a product or object can be compared. Due to the wide range of usage, it became a standard tool in marketing, advertising, and attitude research. Applying this knowledge to the field of emotion research, (Ertl, 1965) and (Mehrabian and Russell, 1974) showed that also emotional adjectives can be reduced to a three-dimensional (affective) space with the dimensions valence/pleasure, arousal, and dominance/stance, see Figure 3a. Thereby, the three dimensions found by (Osgood, 1957) can easily be transformed into the dimensions found by (Mehrabian and Russell, 1974).

Valence can be interpreted as the evaluation dimension, arousal as the activity dimension, and potency can be referred to as the dominance dimension. From his dimensional paradigm (Mehrabian & Russell, 1974), Mehrabian developed a test system called Pleasure-Arousal-Dominance (PAD) emotion model (Mehrabian, 1998). The use of Mehrabian's PAD-test to measure affective experiences represents a generic way to gather self-report-based user data regarding emotions. In the following paragraph the PAD-test is described in more detail.

2.3 The PAD-Emotion Model by Mehrabian

The test is divided into three scales: a 16-item pleasure-displeasure scale, a 9-item arousal-non-arousal scale, and a 9-item dominance-submissiveness scale. The items of the PAD-test are also in the format of a semantic differential. The duration of this test is approximately 7 minutes (Mehrabian, 2007). Alternatively, a short 12-item version exists. Each emotion expressing robot can, thus, be rated in 2-3 minutes. The reliability (internal consistency) of

the full-length version is (α_c : Cronbach's alpha) $\alpha_c=0.97$ for the pleasure scale, $\alpha_c=0.89$ for the arousal scale, and $\alpha_c=0.80$ for the dominance scale. The internal consistency for the abbreviated version is $\alpha_c=0.95$ for pleasure scale, $\alpha_c=0.83$ for the arousal scale, and $\alpha_c=0.78$ for the dominance scale (Mehrabian & Russell, 1974).

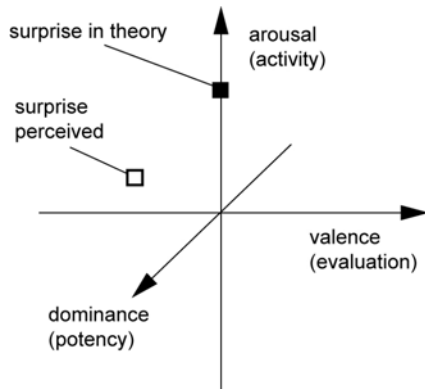


Figure 3a. Results of the semantic differential

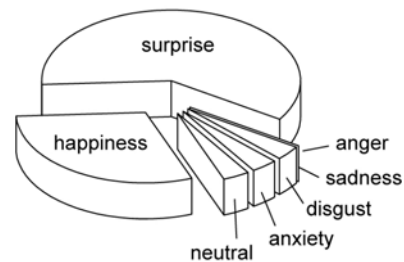


Figure 3b. Results of a categorical test

As a result of the fact that the affective space (with the dimensions valence, arousal and dominance) (Mehrabian & Russell, 1974) or the derived two-dimensional version of the circumplex model of emotion (Russell, 1997) are a common theoretical basis for building emotion expressing robots, only the PAD-model has been introduced as a method to obtain data from affective experiences. There are several other tests measuring attitudes and feelings, e.g. the PANAS (Watson et al., 1988), which, however, are not as easily transferable into the theoretical model of emotion expressing robots.

2.4 Advantages of Dimensional Approaches as the Semantic Differential

Generally, two approaches to evaluate emotion expressing robots exist. On the one hand, dimensional approaches like the semantic differential can be used representing emotion states in continuous affective space. On the other hand, emotions can be treated as quasi-independent categories by asking the subjects to choose the perceived emotion state of the robot from a list of possible emotions (Breazeal, 2002b).

In this chapter, dimensional approaches like the PAD-model are proposed as a generic tool for dimensionally designed emotion expressing robots due to the following advantages:

2.4.1 Test-Theory

A test is conducted in order to examine the performance of the emotion expressing robot. Thus, the test also examines the theory in which the robot is embedded. In order to do so, the test has to be derived from the same theoretical framework as the robot. So, if the robot is embedded in the framework of the circumplex model of Russell (Russell, 1997) or the affective space of Mehrabian (Mehrabian & Russell, 1974) then the test has to be embedded in the same framework. In case of the affective space, which is widely used as a starting point for development, the PAD-test meets this requirement.

2.4.2 Test-Construction

As discussed above, there has to be a substantiated description of the conjunction between latent variable and answering behavior in the test. Since the PAD-model is based on the established and well-grounded model of the semantic space (Osgood et al., 1957), it meets this premise.

2.4.3 Guidelines for Improvements

The result of an evaluation experiment should provide concrete instructions how the evaluated device can be improved. If the robot is theoretically embedded in affective space the semantic differential provides well interpretable data on the quality of the emotion expression on all three dimensions. Figure 3a shows an example with fictive data for the perceived emotion 'surprise' and the theoretical position of 'surprise' in affective space. Due to the fact that certain activation units are linked directly to certain dimensions/areas of the affective space and, moreover, the activation units are linked to certain artificial muscles of the emotion expressing robot, it is possible to conclude the contribution of the artificial facial muscle from the measured position in affective space. Thus, by using the semantic differential more of the gathered information can be processed and interpreted.

Contrarily, if a list of categorical emotions is used for evaluation, only data with "correct hits" is gathered. Such a categorical test would only state that the emotion 'surprise' is identified correctly by a certain percentage. Furthermore, from this result it would be unclear which affective dimension of the displayed emotion is identified well or poorly. In consequence, no conclusion on how to adjust the corresponding joint positions in order to improve the displayed emotion can be drawn. Figure 3b exemplarily shows the fictive result of a categorical evaluation method. It can be noted that about 50% of the participants evaluate the facial expression of the robot as 'surprise', around 30% perceive it as 'happiness' and the rest as other emotions. To this point, no method exists to derive guidelines for improvement of the robot from this evaluation.

2.4.4 Weighting Algorithm

In (Breazeal, 2002b) it is argued that if ten items in a multiple choice test are provided then the probability of chance for each item to be picked is ten percent and, thus, the expected chance probability of each emotion to be picked in such a test would be also ten percent. However, this assumption does not hold. Due to the fact that some emotions share more activation units than others (Grammer & Oberzaucher, 2006) and some lie closer together in the affective space, the probability of each emotion to be picked has to be weighted accounting for these issues. Thus, the expected probability of ten percent of the mentioned example would be increased for similar emotions and decreased for dissimilar emotions. Yet, an algorithm for weighting the expected probability has not been developed. Such an algorithm, however, would be needed for a categorical test since these expected probabilities are used in statistical tests to analyze whether a displayed emotion is classified sufficiently correctly.

2.4.5 Reliability and Validity

Each test, which is used should provide data on its quality. Otherwise, the quality of the result of the test cannot be assessed. The value of a test can be rated by quality measures of classical test theory: reliability, validity, and objectivity. These have been evaluated for dimensional tests, e.g. for the PAD-test. Furthermore, this test has been used in other

studies, e.g. (Valdez & Mehrabian, 2006; Mehrabian et al., 1997), and evaluated towards other tests, which also test affective experiences (Mehrabian, 1997).

2.4.6 Expressive Features of Arbitrary Design

Emotions of biological entities are expressed by a variety of expressive elements from facial muscles over limb motions to color adjustments and acoustic signals. To date, only a framework for the integration of facial elements for expression of emotional states of artificial entities exists (e.g. FACS). However, this framework assumes actuators, which correspond exactly to those of the biological paradigm. If these are not matched well, then the controlled expressive element provides a different contribution to the emotion state to be expressed and, thus, a different emotion will be perceived as is intended. A framework for adjustments of expressive element motions in order to match the intended emotion state better has not been established yet. Dimensional approaches like the PAD-model can be utilized not only to accomplish the desired adjustment, but also to evaluate and adjust the impact of expressive features of arbitrary design (e.g. animal-like) due to the provided guidelines for design improvement (Bittermann et al., 2007).

3. Application of the Guidelines

3.1 Comparative Evaluation Study

In order to illustrate the advantages of dimensional evaluation approaches as the semantic differential for evaluation of emotion expressing robots, two user studies have been conducted exemplarily evaluating the robot head EDDIE, which has been developed at the authors' lab (Sosnowski et al., 2006a). These studies are only intended to show up the differences of both approaches and do not account for demographical issues of the test participants.

The first study uses a test based on a dimensional model of emotions and evaluates the performance of the emotional facial expressions. The second study uses a test based on a categorical emotion model evaluating the same robot. In the following, the system setup is presented in brief, the studies are described, and a discussion is given.

3.1.1 Face Robot EDDIE

EDDIE is a robot head designed for displaying facial expressions, particularly, emotional expressions realizing 13 of the 21 action units of FACS relevant to emotional expressions. In addition to the facial elements, animal-like features, the crown of a cockatoo and the ears of a dragon lizard with special folding mechanisms, are integrated.

EDDIE is encapsulated accepting commands from a higher-level decision and control unit via a serial communication protocol. The desired displayed emotion state can be transmitted based on the three-dimensional affective space representation and feedback is given in affective and joint space representations. An embedded controller manages the transformation between affective space, action units, and joint space. More details on design and control of EDDIE can be found in (Sosnowski et al., 2006a).

3.1.2 Categorical Evaluation Study

A study with 30 participants (researchers of Technische Universität and the University of the Armed Forces, München, 11 female, 19 male, age-mean 30 years) has been conducted.

Six basic emotions have been presented to each individual. The subjects' ratings of the displayed emotions were rated with a multiple-choice test with a seven-item-scale. The result of this study can be seen in Figure 4. On the abscissa the six displayed emotions are shown. For each displayed emotion the amount of assumed emotions by the subjects is presented. For example, 90% of the participants agree in seeing a sad face if a sad face is displayed. For the displayed emotions 'anger' and 'anxiety' about 50% of the answers were correct, the remaining 50% are shared between the other emotions. Evaluating these results, a scientist should be able to draw conclusions in order to improve the robot. Yet, no framework exists in order to formulate new guidelines for robot improvement considering the incorrect answers. Furthermore, it is questionable not to use the data of the incorrect answers as information would be disregarded.

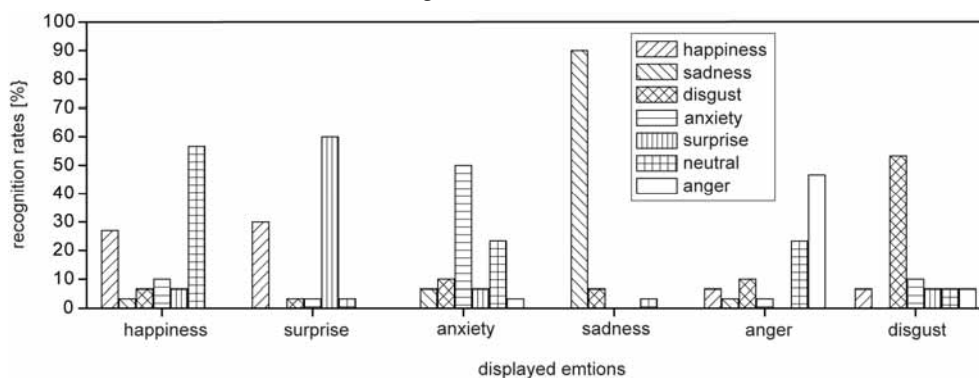


Figure 4. Results of a categorical evaluation study on EDDIE

3.1.3 Dimensional Evaluation Study

A study with 30 participants (students and researchers of the Ludwig-Maximilians-Universität, Munich, 15 female, 15 male, age-mean 25 years) has been conducted. A number of 30 different facial expressions corresponding to certain emotion states has been presented to each subject separately in random order. The subjects' impressions of the shown emotions have been acquired by using a German translation of the semantic differential of Mehrabian. The results of the study are presented in Figure 5 showing the expected emotion values (ground truth) and the values obtained from the study (measurements) in affective space (dimension dominance is not displayed). The results of the study clearly show how each perceived emotion is empirically located in affective space.

3.1.4 Discussion

From these results and the knowledge of the action units actually needed for each specific emotion (Grammer & Oberzaucher, 2006) the quality of the realization of each action unit in the mechatronical robot face can be concluded. Additionally, steps for improvement can be derived from the results. For example, the results in Figure 5 show that the displayed emotion 'fear' has been perceived as a nearly neutral emotion with small values on each dimension. By analyzing this problem, one can compare the amount of action units needed for the intended displayed emotion and the amount of action units realized in the robot. Fear consists of action units 1, 2, 4, 20 and 26. In EDDIE the action units 1 and 2 are

combined for technical reasons and cannot work separately. Furthermore, the action unit 4 was not implemented in the display. Yet, action unit 4 reflects the brow lowerer (*musculus corrugator supercilii*) and is important for the emotion 'fear'. Furthermore, the range of action unit 20 (lip stretcher) could have been too small to show a motion, which people expect from their experience with real human emotions. Based on these conclusions direct tasks of improvement can now easily be derived.

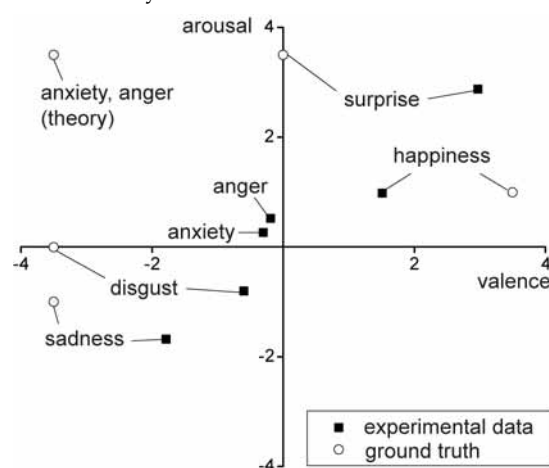


Figure 5. Results of dimensional evaluation study on EDDIE based on the semantic differential approach (without ears and comb)

Not only single emotions but also global deviations of the robot can be assessed. In Figure 5 a global “positive valence shift” can be noted. By comparing the theoretical values of the emotions and the ones actually found, it is obvious that the displayed emotions are all shifted by one or two units to the positive side of the valence dimension. A possible reason for this shift can be noted in Figure 1. The lips of EDDIE are designed in such a way that it seems to smile slightly in most displayed emotion states. Guidelines for improvement are, thus, clear from this point: the lips have to be redesigned. This example shows how even the summarized results can be used to provide new insight into the overall quality of the robot. This is a clear advantage of this method. Taking all the different aspects into consideration, it can be stated that dimensional evaluation methods as the semantic differential approach provide a powerful tool for evaluation of expressive robots by backprojection of joint space into affective space via human perception.

3.2 Integration of Animal-like Features

In a more general context the semantic differential approach can be utilized as a generic means to evaluate the influence of actuated expressive elements of an arbitrary kind on the perceived emotion. The knowledge gained from such evaluation procedures can then be used for the derivation of control commands for those expressive elements. Thereby, actuated expressive features of arbitrary design can be systematically controlled in order to intensify or attenuate the displayed emotion in a particular selected dimension, e.g. valence, arousal, dominance in case of the affective space.

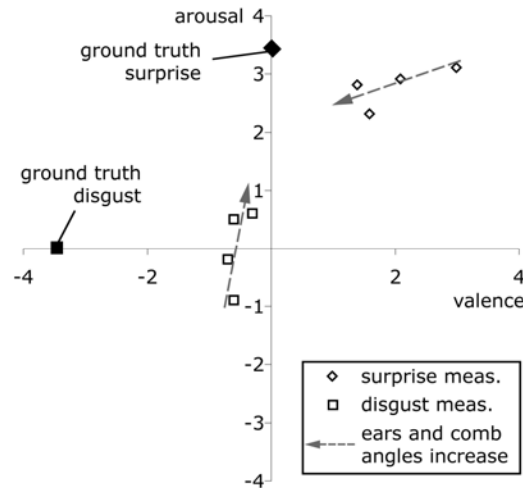


Figure 6. Results for displayed 'surprise' and 'disgust' of a dimensional evaluation study of EDDIE with varied angles of the ears and the comb

3.2.1 A Pilot Study

This is exemplarily shown in an experimental pilot study with 30 participants (15 females, 15 males) evaluating the influence of two animal-like features (crown of a cockatoo and ears of a dragon lizard). In a 2x2 ANOVA design with repeated measures (1. Factor: crown, 2. Factor: ears) it is analyzed whether these two factors shift the observed six basic emotions in affective space. Each factor is realized in four conditions (from fully stilted to dismantled). All six basic emotions are displayed with each combination of the two factors. Afterwards, the participants have to rate each displayed emotion on the verbal semantic differential scale. Every subject has participated in one third of the 96 possible combinations. All data is tested with a Mauchly-test for sphericity. All values are greater than 0.1. Thus, no Greenhouse-Geisser correction is necessary. Missing values are substituted by linear interpolation. Due to incorrect answering behavior some data has to be excluded. For that reason no F-Test can be calculated for 'joy' and 'surprise' (dimension dominance) and 'fear' and 'sadness' (dimensions arousal, dominance). The significant results of the ANOVA are shown in Table 1.

The results suggest that the ears and the crown may have an influence, in particular, for the emotions joy, surprise, anger and disgust. As can be seen in Table 1, mostly the interaction effect between crown and ears becomes significant. For the emotions anger and disgust the animal-like features effect the evaluation of the subjects on all dimensions in affective space. Interestingly, the additional features have a different effect on the evaluation depending on the emotion the robot is expressing. This can also be seen in Figure 6 clearly showing the tendency of the propagation of selected emotion states in affective space while adjusting the attitude angle of the ears and crown simultaneously. While for some emotions a shift towards positive valence and higher arousal side can be noted (e.g. for 'disgust'), the impact of ears and crown is rather reverse for other emotions (e.g. for 'surprise'). The results of the pilot study, thus, suggest further investigations of the usage of new non-humanlike features in emotion expressing robots.

Evaluating the propagation of the perceived emotion 'surprise' in Figure 6 a straightforward control command in order to improve the expression of 'surprise' using the ears and crown in addition to the facial elements is obvious. If ears and crown are fully extended then the perceived emotion 'surprise' moves towards the ground truth. Considering again the global positive shift due to the lip design it can be expected that ground truth 'surprise' is nearly perfectly achievable after a lip redesign using the animal-like features ears and crown. This is a nice result considering that robot and human facial action units differ substantially. The use of additional features, thus, contributes to improve the quality of the artificial emotion expression.

It has successfully been shown that arbitrary expressive features can easily be integrated in a control concept of emotion expressing robots using the semantic differential approach.

emotion	dimension	factor	F-value	p-value
1, 2	V	comb	F(3,12)=4.013	0.034
1, 2	V	ears * comb	F(9,36)=3.631	0.003
1, 2	A	ears * comb	F(9,54)=3.258	0.003
3, 4	V	ears * comb	F(9,18)=5.843	0.001
5, 6	V	ears	F(3,6)=4.835	0.048
5, 6	V	ears * comb	F(9,18)=4.132	0.005
5, 6	A	ears	F(3,6)=67.582	0.000
5, 6	A	comb	F(3,6)=11.987	0.006
5, 6	D	ears	F(3,62)=46.724	0.000
5, 6	D	ears * comb	F(9,18)=9.463	0.000

Table 1. Results of 2x2 ANOVA, repeated measures. (1: happiness, 2: surprise, 3: anxiety, 4: sadness, 5: anger, 6: disgust; V: valence, A: arousal, D: dominance)

4. Conclusions

In this chapter the use of dimensional approaches for evaluation of an emotion expressing robots is discussed. The advantages of dimensional evaluation studies are pointed out and exemplarily illustrated in evaluation studies on face robot EDDIE using a categorical test and the PAD test based on a semantic differential approach. Main arguments supporting the use dimensional models are provided guidelines for design improvement obtained from the evaluation results and a possible integration of expressive elements of arbitrary design. The impact of additional expressive features of EDDIE (ears and comb), which are not covered by a common design framework as, e.g. the Facial Action Coding System, is exemplarily evaluated in a pilot study showing a significant influence of these features on most emotion states. Based on these results the derivation of control commands for integrating these expressive features in a general framework is discussed.

Future work will cover the improvement of the performance of EDDIE by conducting comprehensive studies accounting also for demographical issues and integrating the additional expressive elements based on the framework established in this chapter.

5. Acknowledgments

The authors would like to thank Dr. Albert Mehrabian (University of California, Los Angeles) for inspiring discussions.

This work is supported in part within the DFG excellence initiative research cluster *Cognition for Technical Systems – CoTeSys*, see also www.cotesys.org.

6. References

- Bittermann, A., Kühnlenz, K. & Buss, M. (2007). On the Evaluation of Emotion Expressing Robots, *Proceedings of the IEEE/RAS International Conference on Robotics and Automation (ICRA)*, pp. 2138-2143, 2007, Rome, Italy
- Breazeal, C. L. (2002a). *Designing Sociable Robots*, MIT Press, Cambridge, MA, USA
- Breazeal, C. L. (2002b). Emotion and Sociable Humanoid Robots, *International Journal of Human-Computer Studies*, Vol. 59, No. 1-2, 2002, pp. 119-155
- Blow, M., Dautenhahn, K., Appleby, A., Nehaniv, C. L. & Lee, D. C. (2006). Perception of Robot Smiles and Dimensions for Human-Robot Interaction Design, *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 469-474, 2006, Hatfield, UK
- Ekman, P. (1982). Methods for measuring facial action, In: *Handbook of methods in nonverbal behavior research*, Scherer, K. R. & Ekman, P., (Eds.), pp. 44-90, Cambridge University Press, Cambridge and New York, UK
- Ekman, P. & Friesen, W. V. (1977). *Facial Action Coding*, Consulting Psychologists Press, Palo Alto, CA, USA
- Ertl, S. (1965). Standardization of a semantic differential, *Zeitschrift für Experimentelle und Angewandte Psychologie*, Vol. 12, No. 1, 1965, pp. 22-58
- Grammer, K. & Oberzaucher, E. (2006). The reconstruction of facial expressions in embodied systems: New approaches to an old problem, *ZfJ Mitteilungen*, Vol. 2, 2006, pp. 14-31
- Hamm, A. O. & Vaitl, D. (1993). Emotional induction by visual cues, *Psychologische Rundschau*, Vol. 44, 1993, pp. 143-161
- Hara, F. & Kobayashi, H. (1996). A Face Robot Able to Recognize and Produce Facial Expression, *Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS)*, pp. 1600-1607, 1996, Osaka, Japan
- Marx, W. (1997). Semantische Dimensionen des Wortfeldes der Gefühlsbegriffe, *Zeitschrift für Experimentelle Psychologie*, Vol. 44, 1997, pp. 478-494
- Mehrabian, A. (1997). Comparison of the PAD and PANAS as models for describing emotions and for differentiating anxiety from depression, *Journal of Psychopathology and Behavioral Assessment*, Vol. 19, 1997, 331-357
- Mehrabian, A. (1998). *Manual for a comprehensive system of measures of emotional states: The PAD model*, (Available from Albert Mehrabian, 1130 Alta Mesa Road, Monterey, CA, USA 93940)
- Mehrabian, A. (2007). *The PAD Comprehensive Emotion (Affect, Feeling) Tests*, <http://www.kaaj.com/psych/scales/emotion.html>
- Mehrabian, A. & Russell, J. A. (1974). *An approach to environmental psychology*, MIT Press, Cambridge, MA, USA

- Mehrabian, A., Wihardja, C. & Lyunggren, E. (1997). Emotional correlates of preferences for situation-activity combinations in everyday life, *Genetic, Social, and General Psychology Monographs*, Vol. 123, 1997, pp. 461-477
- Minato, T., Shimada, M., Ishiguro, H. & Itakura, S. (2004). Development of an Android Robot for Studying Human-Robot Interaction, *Proceedings of the International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems (IEA/AIE)*, pp. 424-434, 2004
- Osgood, C. E., Suci, G. J. & Tannenbaum, T. H. (1957). *The measurement of meaning*, University of Illinois Press, Oxford, UK
- Rost, J. (1996). *Lehrbuch Testtheorie, Testkonstruktion*, 1. Aufl., Huber, Bern, Austria
- Russell, J. A. (1997). Reading emotions from and into faces: Resurrecting a dimensional-contextual perspective, In: *The Psychology of Facial Expression*, Russell, J. A. & Fernandes-Dols, J., (Eds.), pp. 295-320, Cambridge University Press, Cambridge, UK
- Sosnowski, S., Kühnlenz, K. & Buss, M. (2006a). EDDIE - An Emotion-Display with Dynamic Intuitive Expressions, *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 569-574, 2006, Hatfield, UK
- Sosnowski, S., Kühnlenz, K., Bittermann, A. & Buss, M. (2006b). Design and Evaluation of Emotion-Display EDDIE, *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2006, Beijing, China
- Valdez, P. & Mehrabian, A. (1994). Effects of color on emotions, *Journal of Experimental Psychology*, Vol. 123, 1994, pp. 394-409
- Watson, D., Clark, L. A. & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales, *Journal of Personality and Social Psychology*, Vol. 54, 1988, pp. 1063-1070
- Zecca, M., Roccella, S., Carrozza, M. C., Miwa, H., Itoh, K., Cappiello, G., Cabibihan, J.-J., Matsumoto, M., Takanobu, H., Dario, P. & Takanishi, A. (2004). On the Development of the Emotion Expression Humanoid Robot WE-4RII with RCH-1, *Proceedings of the IEEE International Conference on Humanoid Robots (Humanoids)*, pp. 235-252, 2004, Los Angeles, CA, USA