

On the Trail of Facial Processing in Autism Spectrum Disorders

Diana Arellano^{1(✉)}, Ulrich Max Schaller², Reinhold Rauh², Volker Helzle¹,
Marc Spicker³, and Oliver Deussen³

¹ Filmakademie Baden-Württemberg, Ludwigsburg, Germany
{diana.arellano,volker.helzle}@filmakademie.de

² University Medical Center Freiburg, Freiburg Im Breisgau, Germany
{reinhold.rauh,ulrich.schaller}@uniklinik-freiburg.de

³ University of Konstanz, Konstanz, Germany
{marc.spicker,oliver.deussen}@uni-konstanz.de
<http://research.animationsinstitut.de/projects/sara/>

Abstract. To investigate the difficulties in communication, socialization abilities, and emotion perception deficits in individuals with autism spectrum disorders (ASD), we propose the project SARA. Its main goal is to assess how abstracted emotional facial expressions influence the categorization of the emotions by children and adolescents with high-functioning ASD. This paper focuses on the first pilot study where an adapted version of the Dynamic Emotional Categorization Test (DECT) was implemented. The results support the validity of the emotional facial animations, which is confirmed by the low performance of subjects with high-functioning ASD when categorizing emotions, in comparison with their peers without ASD.

Keywords: Autism spectrum disorders · Real-time NPR · Abstraction · 3D characters · Dynamic facial animations

1 Introduction

A smooth interaction in social contexts depends in great measure on an adequate internal mental representation of the emotional state of those we interact with. However, people diagnosed with autism spectrum disorders (ASD) present difficulties in social interaction and lesser communication abilities, which are in part the result of emotion perception deficits [1].

To investigate this issue we propose SARA (Stylized Animations for Research in Autism), an ongoing project that studies how abstracted emotional facial animations are categorized by children and adolescents with high-functioning ASD. To carry out our research, we implemented an adapted version of the Dynamic Emotional Categorization Test (DECT) [2], named R-DECT, where dynamic facial animations of physically-plausible 3D characters are generated in real-time, and can be parameterized in terms of speed and intensity.

In this paper we present the results of our first study, the goal of which was to obtain a better understanding of the perceptual and conceptual facial

processing of individuals with and without ASD. To that end, we conducted a feasibility and validation study of the R-DECT to set the basis for the inclusion of abstractions, which constitutes one of the novelties in our project.

2 Background

Impairments in the ability to perceive and categorize facial expressions have been found in a number of psychiatric disorders such as Affective Disorder, Attention-Deficit/ Hyperactivity Disorder (ADHD), Anxiety Disorders, Eating Disorders [3], and Autism Spectrum Disorders (ASD) [4, 5]. Due to the core symptomatology of these disorders, deficits in emotion recognition are multifaceted and show varying degrees of severity.

In 2008 it was estimated that one in 88 children aged 8 years had ASD [6]. However, in 2010 the reports indicated one in 68 children [7]. A direct consequence of this increase has been a growing awareness of the biopsychosocial functioning and of the quality of life of individuals with ASD.

Different fields have conducted exhaustive investigations to understand ASD, proposing theories like “weak central coherence” (WCC), “executive dysfunctioning”, or a “delayed development of theory-of-mind” [8]. A very detailed review of methodologies to develop and train the social skills of people with autism is offered in [9]. One method, especially when working with children and adolescents, consists of the interaction with virtual characters in computer-based applications. The advantages include the possibility of practicing skills in a repetitive way that is more controllable, less threatening and less socially demanding than when interacting with real people. The design of the virtual characters has ranged from very cartoony styles like the ones used in ECHOES VE [10], JeStiMuLE [11], Lifeisgame [12], Photogoo [13], to realistic-human-like styles as seen in [14–16], and going through more artistic versions like the thin-line drawn 2D pedagogical agent developed by [17].

3 The SARA Project

The ongoing research project SARA has been structured to allow synergy between clinical psychology, animation and real-time graphics. The main goal is to better understand the deficits in social cognition in individuals with ASD by means of real-time dynamic facial animations, abstraction techniques, eye tracking techniques, psychophysiological and behavioral measurements.

A number of works demonstrate that these deficits might be related to the amount of details conveyed by facial expressions, as well as to different types of processing of the emotional stimuli, i.e. atypical mutual gaze and eye contact [18–20]. Therefore, one of the main novelties of the project is the use of Non-Photorealistic Rendering (NPR) to abstract the faces of our virtual characters, given that NPR helps in reducing the information load and conveying the presented information more efficiently [21, 22]. Moreover, it provides a painterly representation of the characters, bringing artistic elements into play, which have shown to be of great use in therapies for individuals with ASD [23, 24].

4 DECT

The Dynamic Emotional Categorization Test (DECT) [2] was conceived to assess the feasibility of using real-time animations by comparing virtual characters with human actors. In its very first version, it contained material of two human actors and two virtual characters displaying dynamic facial expressions of the six basic emotions (anger, disgust, fear, happiness, sadness, and surprise) on three intensity levels (weak, medium, and strong). Subjects had to categorize the presented video clips of real actors and animations of virtual characters in a six-alternative forced choice task with the six basic emotions as options (Fig. 1).

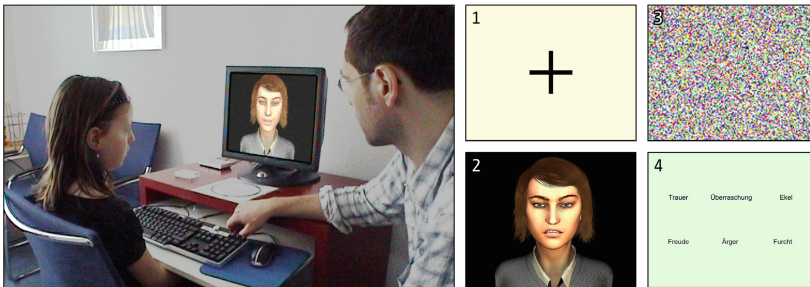


Fig. 1. Setup of the experiment and overview of a DECT trial: (1) Fixation cross that appears when the trial begins, (2) Facial Expression of one of the characters with certain speed and intensity, (3) White-Noise image to load participants' iconic memory with task-irrelevant information, (4) Emotion names as answers options.

The obtained results showed that the three levels of intensity were better reflected in the artificial actors than in the human actors. It was also concluded that using the Agent Framework [25], the software platform to generate the real-time facial animations, the experimenters could take advantage of the manipulability of the facial animations, and achieve desirable features like tailored testing.

In SARA the DECT has been adapted by improving the existent facial animations, integrating real-time NPR algorithms and adding communication with the open source software PsychoPy [26] and an eye tracker RED250 from SMI, which provides a high-speed sampling rate of 250 Hz, and thus high accuracy.

The first study of our project was the Rapid Social Cognition study, which studied the “rapid” aspect of social cognition. For that purpose, we used a version of the DECT that we called R-DECT, which novelty resides in the random presentation of real-time dynamic animations of facial expressions with different speeds (from normal to very fast) and intensities (weak, medium, strong).

The outcome of this study assesses the quality and the validity of the revised real-time and dynamic animations via accuracy rates. This in turn might suggest visual cues or facial features that play a key role in the understanding of perception and categorization of the expressions.

4.1 Virtual Characters

For the pilot study we used two characters: a young female (Nikita) and an older male (Hank) (Fig. 2). Both characters are provided and distributed with the Agent Framework under the Creative Commons Attribution-Non Commercial-ShareAlike 3.0 Unported License.



Fig. 2. Characters used in the R-DECT.

We decided to create human-like realistic characters because we aimed for visual similarity with real actors. Regarding the reduction of the level of detail in faces, it is more plausible to apply NPR techniques (e.g. sketchy or line drawing) to realistic characters and obtain an abstracted representation, than to reconstruct realistic representations from already abstracted characters.

Another important issue to take into consideration was the interactivity in the animations. On the one hand, we are interested in studying the timing in emotion recognition with dynamically moving faces which have higher ecological validity than static photos [5]. Thus our need to have real-time facial animations that can be parameterized in terms of speed and intensity. On the other hand, we plan more complex experiments where interactivity is a key component. In those, we will track the participants' eye gaze, which will originate the synchronized movement of the characters' head and gaze. All this could not be achieved using existing corpora of static images or video recordings of actors, where there is no real-time interactivity.

The animations of the character's facial expressions were produced using the Facial Animation Toolset (FAT)¹ [27], which permits the creation of believable non-linear facial deformations, based on the Facial Action Coding System (FACS) [28]. The FAT system also considers asymmetrical movements of Action Units (AU), allowing the generation of animations where one side of the face presents greater movement intensity than the other, and also in the upper and lower lip area. As a result, the animations are described in terms of FACS AUs and can be manipulated in Frapper's interface. Any value entered for an AU animation controls the intensity corresponding to the set of facial feature points involved.

¹ <http://research.animationsinstitut.de/facial-research-tools/facial-animation-toolset-2015/toolset-download/>.

5 Experiment

In order to assess the feasibility (also with clinical samples), validity and improvement of the animations quality, we tested 39 adolescents that fulfilled the following inclusion criteria: age between 14.0 to 17.9 years and $IQ \geq 70$. The group of neurotypically developed adolescents (NTD group: $n=22$) consisted of 18 male and 4 female participants without any indication of a psychiatric disorder. The group of individuals with high-functioning autism spectrum disorder (ASD group: $n=17$) consisted of 12 male and 5 female participants (Table 1). Diagnosis of ASD (ICD-10: F84.0, F84.1 or F84.5) was established by the international “gold standard” diagnostic procedure by applying the ADOS [29] and ADI-R [30]. Participants received a voucher for a cinema visit or for a book, and also got travel reimbursements in case of arising expenses. The study was approved by the local ethics committee.

Table 1. Sample characteristics (NTD = Neurotypical Development, ASD = Autism Spectrum Disorder, SD = Standard Deviation, ratio = male/female).

	NTD		ASD	
	(n=22, ratio=18/4)		(n=17, ratio=12/5)	
	Mean (SD)	Min - Max	Mean (SD)	Min - Max
Age	16.4 (1.1)	14.2 - 17.7	16.1 (1.1)	14.0 - 17.6
IQ	110.0 (13.3)	82 - 146	107.5 (20.6)	75 - 145

The R-DECT consisted of 2 (characters) $\times$ 6 (basic emotions) $\times$ 3 (intensity levels), resulting in 36 animations. Regarding the speed variable, it was assigned according to a certain scheme to each of the 36 animations, ranging from 1 (normal speed) up to 2.25 times of normal speed. In total, six levels were used (1.00, 1.25, 1.50, 1.75, 2.00, 2.25). However, since speed levels had little impact on results, we do not refer to this variable in the results section.

5.1 Procedure

The R-DECT was part of a battery of tests for rapid social cognition and intuitive moral reasoning assessment, comprising two experimental sessions, each lasting between 1.5 to 2 h. The R-DECT took about 15 min to be carried out and was administered as the first of five tests in the corresponding session. Experimental sessions were run individually in a quiet booth in the lab, where only the participant and the experimenter were present.

6 Results

6.1 Group Differences

In total, 62.2 % of the animations were categorized correctly. Accuracy rates for the NTD group was 65.7 % whereas for the ASD group was 57.7 %.

In order to test whether groups categorized animation-based emotions differently, whether animations with certain basic emotions differed from each other, or whether there are interactions between these two factors, a 2×6 MANOVA with repeated measurements was conducted using individual accuracy rates per basic emotion (collapsed over 6 test items [2 characters and three intensity levels] as dependent variables. This or closely related statistical approaches are common in the area of facial emotion recognition research (e.g. [31]).

The 2×6 MANOVA with repeated measurements yielded no significant interaction between group and basic emotion ($F < 1$). However, the two main effects were significant (*basic emotion*: $F(5,33) = 77.63$, $p < .0001$; *group*: $F(1,37) = 5.36$, $p = .026$). This means that the ASD group performed significantly worse than the NTD group, and that there were significant differences between categorization of basic emotions (Fig. 3).

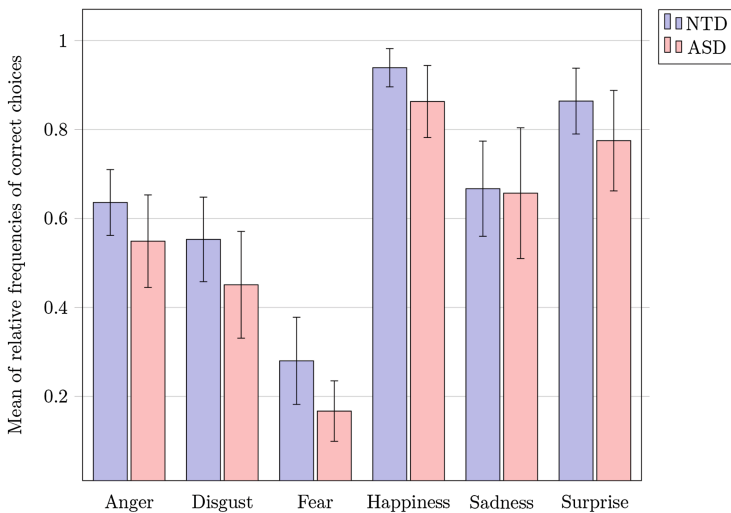


Fig. 3. Mean of relative frequencies of correct categorizations of six basic emotions in adolescents with high-functioning ASD in comparison to neurotypically developed peers (NTD). Error bars represent 95 % CI [confidence interval].

Order of accuracy for basic emotions was the same for both groups and had the following ranking from best to worst: happiness, surprise, sadness, anger, disgust, and fear. Same or similar orders have been found in other studies [31] or in recent meta-analyses [4]. Differences between groups were most pronounced for the basic emotion “fear” (NTD: 28.0 % vs. 16.7 %). However, all post-hoc comparisons did not reach statistical significance.

6.2 Analysis of Materials: Animations Intensities

Concerning the manipulation of the intensity of facial emotions, it is reasonable only to consider the NTD group where typical facial emotion recognition could be

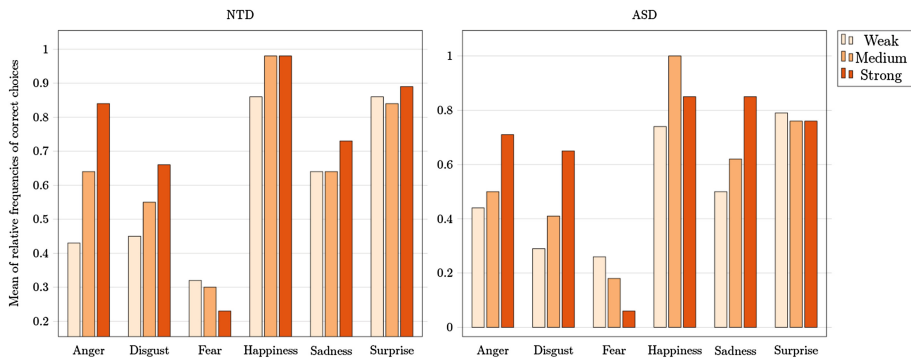


Fig. 4. Mean of relative frequencies of correct categorizations of six basic emotions across differing levels of intensity (weak, medium, strong) in the NTD and ASD group.

	Happiness	Surprise	Anger	Sadness	Disgust	Fear
Happiness	97.7	0.0	0.0	1.1	0.0	1.1
Surprise	4.6	86.4	1.1	0.0	1.14	6.8
Anger	0.0	4.6	73.9	5.7	6.8	9.1
Sadness	0.0	2.3	4.6	68.2	11.4	13.6
Disgust	1.1	0.0	33.0	3.4	60.2	2.3
Fear	23.9	26.1	6.8	1.1	15.9	26.1

Fig. 5. Percentage of chosen emotions (rows) per intended expression (columns) for animations with intensity levels medium or strong for the NTD group.

expected. The results showed that in general, varying intensity from weak over medium to strong affected accuracy rates correspondingly as intended: Weak: 59.5 %, Medium: 65.5 % and Strong: 72.0 %.

Broken down by emotion, most animations followed this pattern, except for “fear” that showed not the intended trend (Fig. 4). A similar pattern was found in the ASD group, where accuracy rates for fear also decreased with stronger intensity.

Confusion Matrices. In order to assess the quality of animations in general and which animations caused miscategorization in particular, confusion matrices between intended emotional expression and chosen expression were computed. A diagrammatic representation of the animations with medium or strong intensity (Fig. 5) shows that most emotional expressions were chosen by the NTD group as intended, with the exception of fear. “Fear” animations were frequently misinterpreted as “surprise”, but surprisingly also as “happiness”.

7 Discussion

We have presented the results of our first pilot study, which constitutes the first study of our project SARA (Stylized Animation for Research in Autism). To that

end, we used an improved version of the DECT (Dynamic Emotional Categorization Test), named R-DECT. This version of the test provided physically-plausible 3D models of virtual characters capable of performing real-time facial expressions of emotions. In addition, the generated facial animations were dynamic in nature, as they could be parameterized according to different levels of speed and intensity.

The R-DECT proved to be suited for assessing the general ability of recognizing emotional facial expressions. The results of the experiment validated the test as the animations should be in a range where no floor effects (nobody recognizes the emotion) or ceiling effects (everybody recognizes the emotions) occur. Moreover, the R-DECT was able to detect the deficits in the ASD group and prove the impact of the different intensities in emotional expressions, as intended, see Figs. 3 and 4.

In order to improve the test, some adjustments of animations have to be made especially in the condition where intensity of the expression was intended to be “medium” or even “strong”. That seems to be the case especially for the animation expressing the emotion “fear”.

The results of these experiments lead to multiple application areas with virtual characters capable of expressing facial emotions. For instance, the parameterizable and real-time aspect of our system represents an advantage and a way to get away from videos and recorded material for teaching emotional facial expressions. It would provide the therapists, psychologists and experimenters freedom from the artists and developers, and the possibility to focus more on the individual differences of the autistic subjects. Hence it is possible to create tailored experiments for processing complex research questions.

As an intervention tool, the DECT could be enhanced to place the animations into another context and to provide immediate feedback about wrong/correct answers, as a sort of supervised learning environment.

Acknowledgments. The SARA project (officially Impact of non-photorealistic rendering for the understanding of emotional facial expressions by children and adolescents with high-functioning Autism Spectrum Disorders) is funded by the DFG - German Research Foundation (AR 892/1-1, DE 620/18-1, RA 764/4-1).

References

1. Hudepohl, M.B., Robins, D.L., King, T.Z., Henrich, C.: The role of emotion perception in adaptive functioning of individuals with high-functioning autism spectrum disorders. In: *Autism: Int. J. Res. Pract.* (2013)
2. Rauh, R., Schaller, U.M.: *Categorical Perception of Emotional Facial Expressions in Video Clips with Natural and Artificial Actors: A Pilot Study*. Technical report ALU-KJPP-2009-001, University of Freiburg (2009)
3. Collin, L., Bindra, J., Raju, M., Gillberg, C., Minnis, H.: Facial emotion recognition in child psychiatry: a systematic review. *Res. Dev. Disabil.* **34**(5), 1505–1520 (2013)
4. Lozier, L.M., Vanmeter, J.W., Marsh, A.A.: Impairments in facial affect recognition associated with autism spectrum disorders: a meta-analysis. *Dev. Psychopathol.* **26**(4 Pt 1), 933–945 (2014)

5. Uljarevic, M., Hamilton, A.: Recognition of emotions in autism: a formal meta-analysis. *J. Autism Dev. Disord.* **43**(7), 1517–1526 (2013)
6. Baio, J.: Prevalence of autism spectrum disorders autism and developmental disabilities monitoring network, 14 sites, United States, 2008. *Surveill. Summ.* **61**(SS03), 1–19 (2012)
7. Baio, J.: Prevalence of autism spectrum disorder among children aged 8 years autism and developmental disabilities monitoring network, 11 sites, United States, 2010. *Surveill. Summ.* **63**(2), 1–21 (2014)
8. Hill, E.L.: Executive dysfunction in autism. *Trends Cogn. Sci.* **8**(1), 2632 (2004)
9. Ospina, M.B., Seida, J.K., Clark, B., Karkhaneh, M., Hartling, L., Vandermeer, B., Smith, V.: Behavioural and developmental interventions for autism spectrum disorder: a clinical systematic review. *Autism* **3**(11), e3755 (2008)
10. Alcorn, A., Pain, H., Rajendran, G., Smith, T., Lemon, O., Porayska-Pomsta, K., Foster, M.E., Avramides, K., Frauenberger, C., Bernardini, S.: Social communication between virtual characters and children with autism. In: Biswas, G., Bull, S., Kay, J., Mitrovic, A. (eds.) *AIED 2011. LNCS*, vol. 6738, pp. 7–14. Springer, Heidelberg (2011)
11. Serret, S., Hun, S., Iakimova, G., Lozada, J., Anastassova, M., Santos, A., Vesperini, S., Askenazy, F.: Facing the challenge of teaching emotions to individuals with low- and high-functioning autism using a new Serious game: a pilot study. *Molecular Autism*. 5(37), (2014)
12. Alves, S., Marques, A., Queirós, C., Orvalho, V.: LIFEisGAME prototype: a serious game about emotions for children with autism spectrum disorders. *PsychNology J.* **11**(3), 191–211 (2013)
13. Hourcade, J.P., Bullock-Rest, N.E., Hansen, T.E.: Multitouch tablet applications and activities to enhance the social skills of children with autism spectrum disorders. *Pers. Ubiquitous Comput.* **16**(2), 157–168 (2012)
14. Whyte, E.M., Smyth, J.M., Scherf, S.: Designing serious game interventions for Individuals with autism. *J. Autism Dev. Disord.* (2014)
15. Baron-Cohen, S., Golan, O., Ashwin, E.: Can emotion recognition be taught to children with autism spectrum conditions? *Phil. Trans. R. Soc. B.* **364**, 3567–74 (2009)
16. Milne, M.K., Luerssen, M.H., Lewis, T.W., Leibbrandt, R.E., Powers, D.M.: Designing and evaluating interactive agents as social skills tutors for children with autism spectrum disorder. In: *Conversational Agents and Natural Language Interaction: Techniques and Effective Practices*, pp. 23–48. IGI Global, USA (2011)
17. Grawemeyer, B., Johnson, H., Brosnan, M., Ashwin, E., Benton, L.: Developing an embodied pedagogical agent with and for young people with autism spectrum disorder. In: Cerri, S.A., Clancey, W.J., Papadourakis, G., Panourgia, K. (eds.) *ITS 2012. LNCS*, vol. 7315, pp. 262–267. Springer, Heidelberg (2012)
18. Kättsyri, J., Saalasti, S., Tiippana, K., von Wendt, L., Sams, M.: Impaired recognition of facial emotions from low-spatial frequencies in Asperger syndrome. *Neuropsychologia*. **46**, 1888–1897 (2008)
19. Valla, J.M., Maendel, J.W., Ganzel, B.L., Barsky, A.R., Belmonte, M.K.: Autistic trait interactions underlie sex-dependent facial recognition abilities in the normal population. *Front Psychol.* 4(286) (2013)
20. Strauss, M.S., Newell, L.C., Best, C.A., Hannigen, S.F., Gastgeb, H.Z., Giovannelli, J.L.: The development of facial gender categorization in individuals with and without autism: the impact of typicality. *J Autism Dev Disord.* **42**(9), 1847–1855 (2012)

21. Götzelmann, T.: Towards non-photorealistic rendering in car navigation. *J. Comput. Eng. Inf. Technol.* **1**(1) (2012)
22. Lee, J., Kim, H., Kim, M., Park, H., Kim, H.: Non-Photorealistic Rendering applied Semantic LOD. In: *HCI KOREA 2015*. pp. 1–6. Hanbit Media, Inc. (2015)
23. Epp, K.M.: Outcome-based evaluation of a social skills program using art therapy and group therapy for children on the autism spectrum. *Child. Sch.* **30**(1), 27–36 (2008)
24. Emery, M.J.: Art therapy as an intervention for autism. *Art Ther.: J. Am. Art Ther. Assoc.* **21**(3), 143–147 (2004)
25. Arellano, D., Helzle, V., Schaller, U.M., Rauh, R.: Animated faces, abstractions and autism. In: Bickmore, T., Marsella, S., Sidner, C. (eds.) *IVA 2014. LNCS*, vol. 8637, pp. 22–25. Springer, Heidelberg (2014)
26. Peirce, J. W.: Generating Stimuli for Neuroscience Using PsychoPy. *Front. Neuroinform.* **2**(10) (2009)
27. Helzle, V., Biehn, C., Schlömer, T., Linner, F.: Adaptable setup for performance driven facial animation. In: *ACM SIGGRAPH 2004 Sketches*, p. 54 (2004)
28. Ekman, P., Friesen, W., Hager, J.: *The Facial Action Coding System*. Weidenfeld & Nicolson, London (2002)
29. Rühl, D., Bölte, S., Feineis-Matthews, S., Poustka, F.: *ADOS: Diagnostische Beobachtungsskala für Autistische Störungen*. Huber, Bern (2004)
30. Bölte, S., Rühl, D., Schmötzer, G., Poustka, F.: *ADI-R: Diagnostisches Interview für Autismus - Revidiert German version of Autism Diagnostic Interview - Revised*. Huber, Bern (2006)
31. Smith, M.J.L., Montagne, B., Perrett, D.I., Gill, M., Gallagher, L.: Detecting subtle facial emotion recognition deficits in high-functioning Autism using dynamic stimuli of varying intensities. *Neuropsychologia* **48**(9), 2777–2781 (2010)