Emotion Recognition in Autism Spectrum Disorder: Does Stylization Help?

Marc Spicker[1] Diana Arellano[2] Ulrich Schaller[3] Reinhold Rauh[3] Volker Helzle[2] Oliver Deussen[1]

[1]University of Konstanz, Germany
[2]Filmakademie Baden-Württemberg, Germany
[3]University of Freiburg, Germany

Figure 1: *The emotion happiness shown by a female virtual character rendered realistic (left) and in different stylized variants.*

Abstract

We investigate the effect that stylized facial expressions have on the perception and categorization of emotions by participants with high-functioning Autism Spectrum Disorder (ASD) in contrast to two control samples: one with Attention-Deficit/Hyperactivity Disorder (ADHD), and one with neurotypically developed peers (NTD). Real-time Non-Photorealistic Rendering (NPR) techniques with different levels of abstraction are applied to stylize two animated virtual characters performing expressions for six basic emotions. Our results show that the accuracy rates of the ASD group were unaffected by the NPR styles and reached about the same performance as for the characters with realistic-looking appearance. This effect, however, was not seen in the ADHD and NTD groups.

Keywords: facial animation, emotion recognition, non-photorealistic rendering, autism spectrum disorder

Concepts: •**Computing methodologies** → **Animation; Non-photorealistic rendering;** •**Applied computing** → **Psychology;**

1 Introduction

The ability to perceive emotions and other affective traits from human faces is considered the result of a "social training" that occurs from early childhood and develops during the adolescence and adulthood. However, individuals with Autism Spectrum Disorders (ASD) present difficulties not only judging someone's emotions from their facial expressions [Kennedy and Adolphs 2012], but also in face processing in general [Harms et al. 2010]. There is little consensus regarding the causes of the impairments [Uljarevic and Hamilton 2013][Lozier et al. 2014], but there are indications that an atypical local-oriented strategy while processing faces might be one reason [Behrmann et al. 2006][Deruelle et al. 2008].

In this paper we present a study that investigates the differences in perception and categorization of abstracted emotional facial expressions in virtual characters between children and adolescents with ASD, and those Neuro-Typically Developed (NTD) and those with Attention-Deficit/Hyperactivity Disorder (ADHD). We decided to use facial expressions with variable amount of details, and therefore varying information load [Kyprianidis and Kang 2011] to assess the impact on emotion recognition due to the atypical face processing in individuals with ASD, in comparison with their NTD and ADHD peers. Thus we formulated two hypotheses:

a) For stimuli with realistic-looking appearance, the NTD group should have higher accuracy (percentage of accurately identified emotions) in recognition rates than the ADHD group, which in turn should be higher than those of the ASD group.

b) For all styles (realistic-looking and NPR rendered stimuli), we presume that stylization and abstraction affects accuracy rates of ASD differently compared to ADHD and NTD, due to an atypical face processing in ASD that should be sensitive to features of the facial stimuli.

For this study, we carried out the NPR-DECT test, an adaptation of the "Dynamic Emotion Categorization Test" (DECT) [Rauh and Schaller 2009]. In contrast to the DECT, which only provided realistic-looking emotional facial expressions, the NPR-DECT relies on Non-Photorealistic Rendering (NPR) algorithms to generate stylized facial expressions. The real-time capabilities give the possibility for interactive parametrization (i.e. speed, intensity, or abstraction styles), allowing the experimenters to manipulate the stimuli. Moreover, it provides possibilities for future interactive applications (e.g. head and gaze animation controlled by the user's gaze data, or stylization methods for real video streams).

The results showed that the NTD group performed clearly better than the ASD group, even when categorizing abstracted emotions. An interesting observation is the lack of preference for any NPR style by participants in the ASD group. This is coherent with other recent research, which showed that different levels of detail in avatars created a roughly equivalent verbal and non-verbal social behavior in children with ASD [Carter et al. 2016].

SAP '16, July 22 - 23, 2016, Anaheim, CA, USA
ISBN: ACM 978-1-4503-4383-1/16/07...$15.00
DOI: http://dx.doi.org/10.1145/2931002.2931004

2 Background

Previous studies have suggested that individuals with ASD process faces and decode facial expressions differently than NTD individuals [Harms et al. 2010]. Two theoretical frameworks have addressed these differences. One is Weak Central Coherence (WCC) that postulates a deficit in global processing and a detail-focused cognitive style [Happé and Frith 2006]. The other is the Enhanced Perceptual Functioning theory (EPF), an alternative to the WCC that re-asserts the principle of locally oriented and low-level perception, without making assumptions about the quality and quantity of global processing [Mottron et al. 2006]. Rosset et al. [2008] compared the processing of emotional expressions of real faces, human cartoon and non-human cartoon faces in children with ASD to two groups of NTD children, finding that those with ASD relied on a local-oriented strategy. This atypical local strategy have also been reported by Behrmann et al. [2006] and Deruelle et al. [2008]. Pelphrey et al. [2002] analyzed visual scan paths of five high-functioning ASD and five NTD adult males as they viewed photographs of human facial expressions. They found out erratic and disorganized scan paths in the ASD participants, and noticed that they spent most of the time viewing non-feature areas of the faces and a smaller percentage of time examining core features such as the nose, mouth, and, especially, the eyes. These findings have also been supported by other works ([Kirchner et al. 2011], [Valla et al. 2013]).

Virtual Characters

A number of studies have confirmed the advantages of using virtual characters (VC) in research on autism. The DECT test [Rauh and Schaller 2009] assessed the feasibility of using real-time animations of realistic-looking virtual characters by contrasting them with videos of human actors performing emotional facial expressions. It showed that the recognition rates using recorded humans or VCs were highly correlated. Tartaro and Cassell [2008] suggested that collaborative narratives with virtual peers may offer a structured setting for using contingent discourse. Georgescu et al. [2014] reviewed previous research done with virtual characters and virtual reality environments (VRE). They concluded that these provide great value for experimental paradigms in social cognition, specifically those related with non-verbal behavior and perception in participants with high-functioning autism, because it allows to grasp the full extent of the social world in a controlled manner.

Stylization

Regarding visual representation, recent studies with autistic children have used VCs as tutors or play companions either with cartoony styles (e.g. [Alcorn et al. 2011], [Alves et al. 2013], [Serret et al. 2014]), or more realistic-looking ones (e.g. [Baron-Cohen et al. 2009], [Milne et al. 2011]). To better understand the differences in style representation and perception of VCs, other studies have endeavored in finding the elements that produced a more appealing, realistic or believable VC. For instance, Grawemeyer et al. [2012] developed a pedagogical agent with the aid of four young people with ASD, resulting in thin line drawn 2D characters. McDonnell et al. [2012] investigated the effects of typical rendering styles in CG productions. Their results showed that participants, all NTDs, were so focused on the given task that the characters appearance was not noticed. Hyde et al. [2013] studied how rendering style and facial motion affect impressions of character personality. They found that slightly exaggerated realistic-looking characters appeared more likeable and intelligent than slightly exaggerated cartoon characters. Zell et al. [2015] observed that shape in stylized characters is the main contributor of perceived realism, whereas material mainly affects perceived appeal, attractiveness and eeriness. Regarding the intensity of facial expressions (anger, happiness, sadness, surprise and neutral), shape was the main factor but material had no significant influence. Mäkäräinen et al. [2015] studied the effects of realism and magnitude of facial expressions, showing that contrary to what was expected, there are cases where stimuli rated high in strangeness produced a positive emotional response. More recently, Carter et al. [2016] found that changing the visual complexity of avatars (video, computer-generated, and cartoon) did not significantly affect any social interaction behaviors of children with ASD.

ADHD

As far as we could observe, there is a gap between studies of facial categorization with participants with ASD in comparison to other clinical groups like ADHD, and studies of facial perception of stylized characters with NTD participants. We included the ADHD group as clinical comparison group because previous studies tend to indicate that these children have difficulties with facial recognition compared to NTDs, particularly in emotion recognition [Collin et al. 2013], and (slightly) better recognition rates than the ASD group [Bora and Pantelis 2015]. The ADHD group also served as control group to confirm that the pattern of impairments are specific to the ASD group, and not due to psychopathology in general.

3 Visual Abstraction

Realistic representations, where the highest level of detail is generally preferred, usually contain more information than necessary to transmit intended information. Therefore, artists typically remove details and use abstraction for a more effective visual communication [Kyprianidis and Kang 2011]. Abstraction is a continuous process for communicating a specific thing to a more general concept, as depicted in Figure 2.

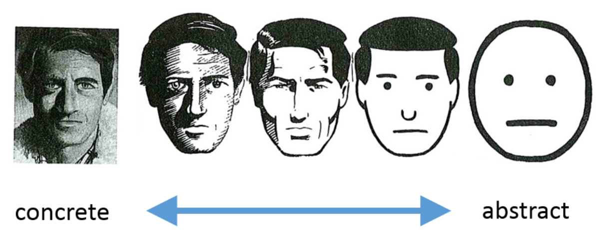

Figure 2: *Different levels of abstraction for the representation of a specific thing (a person), describing an abstract concept (a face and an emotion)[McCloud 1994].*

Non-Photorealistic Rendering (NPR) achieves a variation in the level of abstraction, adapting images to "focus the viewer's attention" [Gooch and Gooch 2001]. It attempts to model the physical properties of an artistic medium such as watercolor, oil, or pen and ink. Therefore, research on NPR aims at the automatic creation of artistic images with a previously defined communication goal. We want to use the abstraction property to reduce the information load when looking at the character's faces to convey the presented information more efficiently. The chosen styles meet the following criteria:

- Large visual distance between each style to cover a large portion of the NPR spectrum.
- Intrinsic possibility to vary the level of abstraction.
- Run at interactive frame rates and are temporally coherent.

Each algorithm was implemented as a multi-pass rendering process in OpenGL to ensure the real-time capability. Table 1 lists the complete parameter set used for each algorithm and abstraction used in the experiments. For a more detailed description of the parameters and their usage, please refer to the original contributors. We used a resolution of 1280×1024 throughout our experiments. Figure 3 shows the virtual characters used in the experiment rendered with a realistic appearance: Hank (elderly male) and Nikita (young female).

3.1 Image Abstraction

This style is inspired by paintings and painting-like images, where absence of fine-grained texture details and increased sharpness of edges are two relevant visual characteristics [Papari et al. 2007]. To reproduce the style, our algorithm is based on Kyprianidis et al. [2009] and Kyprianidis and Kang [2011] which use directional information from the input image to guide a filtering that removes small details but keeps regions boundaries intact. A generalization of the flow-based anisotropy-guided Kuwahara filtering is used to blur similar regions, while leaving local shapes and structures of the input intact. First, a rotational symmetric derivative filter is used to approximate the local structure, the structure tensor. In a next step, these tensors are Gaussian smoothed to avoid artefacts due to noise or very small features. The eigenvectors of the smoothed tensors are used to determine the local orientations and anisotropy. As a last step, an ellipse is defined whose major axes are aligned to the local eigenvectors and scaled according to the eigenvalues. This ellipse is divided into 8 radial regions, and the colors of each region are summed up and then weighted by a Gaussian kernel within the elliptical area. The final output is the summed up results for all 8 regions. Two parameters give the user control over the abstraction of the algorithm. One is the filter radius that controls the size of features to be filtered out: Choosing a large radius results in blurring-out larger features. The second parameter controls the sharpness of the output by using it as exponent in the weighting calculation of the elliptic subregions. Figure 4 shows the image abstraction.

3.2 Coherent Line Drawing (CLD)

Line drawing effectively conveys shapes and outlines to the viewer with simple primitives: lines. Despite the vast amount of computer vision algorithms for line extraction and edge detection, many of them suffer from noisy inputs and produce incoherent outputs, such as line fragments. We use the algorithm of Kang et al. [2007] which creates coherent and artistic-looking lines. The idea behind it, which is closely related to line integral convolution, is to use the directional information of the local image structure to guide an anisotropic difference-of-Gaussian (DoG) filtering. The directional information is extracted from the input image by creating an edge tangent flow field. The DoG-filtering collects filter responses along the flow to determine the local edge strength, creating coherent lines and suppressing noise. Different levels of abstraction are achieved by varying the DoG filter kernel size, the amount of blur during the edge tangent field construction, and the integration length along the flow. Figure 5 shows results of the coherent line drawing algorithm.

Figure 3: *The male (Hank) and female (Nikita) characters used for the experiments in a neutral pose and rendered realistically.*

Figure 4: *Image Abstraction (Left: medium, right: high abstraction) applied on a male and female character.*

3.3 Pencil Drawing

Pencil drawing is one of the most fundamental techniques in visual arts to abstract human perception of natural scenes. By converting input scenes into just lines and shading, a great number of details are removed, while keeping the object boundaries and plasticity of the rendered objects by preserving their shading. We based our implementation on the work of Lu et al. [2012], which combines two common steps in the human drawing process: pencil sketching to express the general structure of the scene, and tone drawing for shading. The method begins by computing the images gradients as a simple forward difference and clustering them according to discrete slopes with eight different angles. The gradient direction with maximal magnitude for each pixel is determined and a kernel with this direction is used to create the pencil drawing line. The tonal value is estimated from the input image by converting it into a grayscale image and applying an empirically obtained heuristic model defined by three components. The tonal value can either be used directly as brightness to create a grayscale image, or as value in the YUV color space for colorized results, providing a simple mean of abstraction. In addition, the length of the created lines in the pencil sketch can be controlled by the directional kernel size. Figure 6 shows the results of the pencil drawing.

	Param.	Medium Abs.	High Abs.
Coherent Line Drawing	r	5	8
	τ	0.98	0.97
	σ_1	1	0.7
	σ_3	4	6
Pencil Drawing	Monochrome	no	no
	Kernel Size	200	120
	W_1 / W_2 / W_3	11 / 37 / 52	11 / 37 / 52
Image Abstraction	r	6	10
	q	6	16
Watercolor	Noise Scale	2.85	2.85
	Noise Intensity	45	80

Table 1: *Parameters used for the rendering algorithms and their abstraction levels.*

Figure 5: *Coherent Line Drawing (Left: medium, right: high abstraction) applied on a male and female character.*

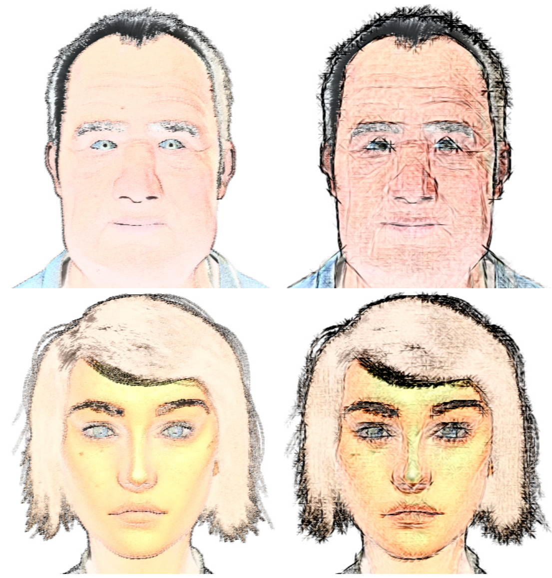

Figure 6: *Pencil Drawing (Left: medium, right: high abstraction) applied on a male and female character.*

3.4 Watercolor

Watercolor painting is an artistic style that creates the effect of water-dissolved colors on paper, or similar surfaces. Two main styles of watercolor painting can be differentiated: wet-on-wet, where colors are added onto not yet dried colors on the medium, and wet-on-dry, where the underlying color is already dry. The prominent features of real watercolor images, such as brilliant colors, subtle variation of color saturation, and visibility of the underlying medium are the result of a complex interaction between pigments, water and the support medium [Bousseau et al. 2006]. We decided to use this style not only for the abstraction it provides, but also because it is used by many individuals with ASD to express themselves through painting, or during therapies that introduce artistic elements [Tataroğlu 2013]. Our algorithm is based on the work of Luft et al. [2006] whose approach simplifies the visual complexity and imitates the natural effects of watercolor. The effect of painting wet-on-wet is imitated by varying the pigment granulation according to the underlying paper structure, affecting the water flowing process. This process was simulated by modifying the input image according to a noise-based displacement vector. The level of abstraction of the watercolor algorithm is achieved by varying the noise intensity. Figure 7 shows the results of the watercolor algorithm.

4 Experiment

A total of 62 participants (all male; ASD is diagnosed about four times more often in males than females according to the American Psychiatric Association [2013]) within an age range between 13.0 and 18.0 years and IQ of at least 85 took part in the experiment. 24 of them were diagnosed with ADHD, 16 with high-functioning ASD and 22 were neurotypically developed (NTD) peers. This last group was recruited from local schools and showed no indication of any psychiatric disorder according to screening questionnaires. All subjects were diagnosed according to ICD-10 criteria for ADHD (F90.0, F90.1) and ASD (F84.0, F84.1, F84.5), respectively. Especially, all patients with ASD were diagnosed based on the "gold standard" with the Autism Diagnostic Observation Schedule (ADOS-G) and with the Autism Diagnostic Interview-Revised (ADI-R). All participants had normal or corrected-to-normal vision and were compensated for their time. The study protocol has been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. To examine the ability of emotion categorization in dynamic facial stimuli, we conducted the NPR-DECT that included stylized representations of the characters. The test was developed in the Frapper framework[1] that provides an interface in which psychologists or experimenters can interactively guide participants through the test sessions. The stimuli consisted of 108 animations compliant with the Facial Action Coding System (FACS) [Ekman et al. 2002], 2 characters (male and female, see Figure 3), 6 basic emotions (anger, disgust, fear, happiness, sadness, and surprise), 5 styles (original, image abstraction, coherent line drawing, pencil drawing, watercolor), and 2 abstraction levels (medium and high). For the original, or realistic-looking style, no abstraction levels exist. The 2 characters were rigged using the Adaptable Facial Setup (AFS) [Helzle et al. 2004], a tool-set that relies on a complete motion capture-based library of deformations, generating high quality, natural and non-linear deformations. Once the characters were finished, they were exported into Frapper. The facial movements are described in terms of Action Units (AUs), providing a parameterizable way of creating animated facial expressions, which could then be translated to any other character with a similar FACS-based rig. In order to keep the number of animations within a reasonable range, intensity of emotions and speed of animation were kept constant across all animations in this experiment. Table 2 gives an overview of each parameter and the associated possible values.

Parameter	#	Values
Character	2	Male, Female
Emotion	6	Anger, Disgust, Fear, Happiness, Sadness, Surprise
Style	5	Original, Image Abstraction, Coherent Line Drawing, Pencil Drawing, Watercolor
Abstraction	2	Medium, High

Table 2: *Parameter space and values used in the experiments.*

The NPR-DECT had an average duration of 25 minutes. One trial consisted of a fixation cross presented for 0.5s, followed by a character's animation displaying a basic emotion rendered in one of the

[1] research.animationsinstitut.de/frapper/

Figure 7: *Watercolors (Left: medium, right: high abstraction) applied on a male and female character.*

distinct styles. Once the animation was finished, a response screen with the six possible basic emotions was shown until the participant pressed a key on the numeric keypad corresponding to an emotion name (e.g. 1-happiness, 2-sadness, 3-anger, 7-fear, 8-disgust and 9-surprise). Trials were presented in a pseudo-randomized order, and the association between numeric keys and emotion names was counterbalanced across participants to avoid order effects. Additionally, the participants were asked to rate the recognizability and likeability of the different combinations of characters, rendering style, and abstraction level on a scale from 1 (very good) to 7 (very bad). In the following we will refer to the different rendering styles by their abbreviations: Original with realistic appearance (Ori.), Image Abstraction (IA), Coherent Line Drawing (CLD), Pencil Drawing (PD), Watercolor (W). For the abstraction levels we will use (M) for medium and (H) for high abstraction.

5 Results

5.1 NPR Style and Level of Abstraction

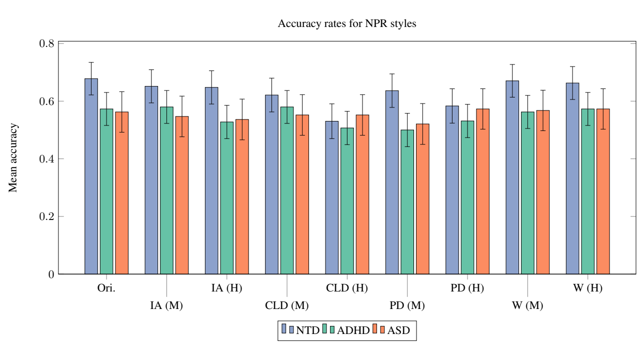

Figure 8: *Accuracy rates for NPR styles with varying abstraction (M: medium, H: high). Aggregated: Emotions, Characters. Error bars represent 95% CI.*

Concerning accuracy rates, the main analysis consisted of a 9x3 MANOVA with repeated measurements with the factors *style parameter* (NPR style and level of abstraction) and *diagnostic group*. The interaction between the factors did not reach significance ($F(16, 104) = 1.24$, $p = .250$), whereas both main effects turned out to be significant (*style parameter*: $F(8, 52) = 3.35$, $p = .004$; *diagnostic group*: $F(2, 59) = 7.84$, $p = .001$). In the follow-up analyses, MANOVAs with repeated measurements for each diagnostic group revealed a significant main effect for *style parameter* for the NTD and ADHD group ($F(8, 14) = 3.28$, $p = .025$; and $F(8, 16) = 2.78$, $p = .039$), but not for the ASD group ($F(8, 8) = 1.14$, $p = .427$). In sum, an "abstracted faces effect" showed up in the NTD and ADHD group, but not in the ASD group, which showed an indifference to the stylizations (Hypothesis 2). Accuracy rates aggregated across emotions and virtual characters are displayed in Figure 8. It can be observed that each NPR rendering style had a lower accuracy rate than the realistic-looking counterpart (Ori.). For the ASD group, accuracy rates were not affected by rendering styles and stay around their total mean of 55.4%. This effect is not seen in the NTD group, and to some extent neither in the ADHD group. It is also interesting to note that increasing the abstraction from medium to high has an overall negative effect on the accuracy for the Image Abstraction and Coherent Line Drawing styles, no effect for the Watercolor style, and a positive effect only for the Pencil Drawing. The accuracy rankings between the medium and the high abstraction remained the same for the NTD group, but changed for the other two groups with no obvious pattern. Between the ADHD and ASD groups the only algorithm that performed well throughout was the Watercolor algorithm. The overall accuracy rates of the NTD group is above those of the two other groups for all styles and abstraction intensities, as formulated in our first hypothesis.

5.2 Emotions

Figure 9 shows accuracy rates for the 6 basic emotions used in the experiments, for each of the diagnostic groups and NPR styles, to investigate if there are any relationships. Characters and abstraction levels are aggregated in this case. Looking at the rank ordering of the emotions, it can be seen that happiness has the highest rank with above 90% accuracy, followed by surprise ($\sim$80%). Sadness, disgust and anger are always in the mid-ranks and change positions between the test groups ($\sim$40-66%). Fear recognition was the lowest in all conditions with an accuracy level that is near chance level ($\sim$10-19%). In particular for the emotion disgust in all rendering styles, the ASD group outperformed the ADHD group. The opposite occurred for sadness in all rendering styles. Leaving out fear due to its low accuracy rates, there is no consistent trend of better performance of one group over another for the remaining emotions, but the expected good performance of the NTD group (first hypothesis) and the overall slightly better performance of ADHD over ASD for the realistic stimuli. Regarding the NPR styles, accuracy ratings for the different emotions show that none of them performed better or worse at conveying a single emotion. Styles that perform well for a group (e.g. Watercolor) do so for different emotions, and the accuracy primarily depends on the shown emotion, indicating that the corresponding facial animation outweighs the importance of the NPR style.

5.3 Characters

Figure 10 depicts the accuracy rates for the male and female characters for the different diagnostic groups. The attributes *NPR style*, *abstraction level* and *emotion* are aggregated. It can be seen that the NTD group has the best performance for both characters (first and second hypothesis), while the other two groups have a similar, but worse performance. For both the male and female character the performance difference is about the same ($\sim$14.5%). Regardless of the groups, using the female character results in an higher accuracy.

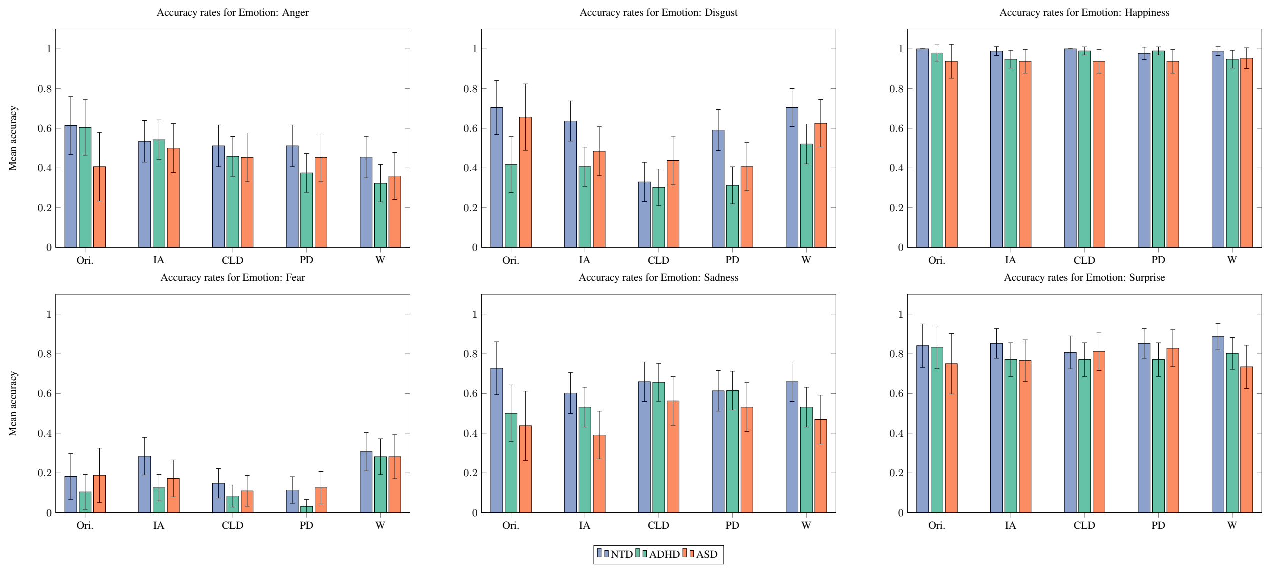

Figure 9: *Accuracy rates for each of the 6 emotions and their respective rendering styles. Aggregated: Characters and abstraction levels. Error bars represent 95% CI.*

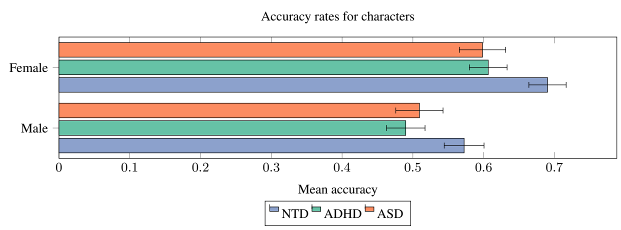

Figure 10: *Accuracy rates for the male and female character depending on test group. Aggregated: Emotions, NPR styles, abstraction levels. Error bars represent 95% CI.*

5.4 Recognizability and Likeability

Table 3 shows the results of participants rating the recognizability and likeability of each NPR style, abstraction level, and virtual character (VC): male (M) and female (F). Ratings vary from 1 (very good) to 7 (very bad) and are color-coded so that worse ratings are displayed with a more intense red color. Concerning recognizability, a 9x2x3 MANOVA with repeated measurements revealed a tendency for the interaction of *style parameter* (NPR style and abstraction level) and *character* ($F(8, 51) = 1.96$, $p = .072$), and a significant main effect of *style parameter* ($F(8, 51) = 12.93$, $p < .0001$). Regarding likeability, the measurements again revealed a significant interaction of *style parameter* and *character* ($F(8, 51) = 2.44$, $p = .026$), and a significant main effect of *style parameter* ($F(8, 51) = 9.20$, $p < .0001$). It indicates that both recognizability and likeability ratings were reliably influenced by the style parameter, and additionally, by the underlying character. To elucidate the impact of different NPR styles and the two levels of abstraction, an additional 4x2x2x3 MANOVA with repeated measurements was conducted. For recognizability, it revealed a significant interaction of *NPR style* and *character* ($F(3, 56) = 3.26$, $p = .028$), and also significant main effects of *NPR style* ($F(3, 56) = 27.06$, $p < .0001$) and *level of abstraction* ($F(1, 58) = 38.57$, $p < .0001$). For likeability, significant interactions were found between *character* and *NPR style* ($F(3, 56) = 3.64$, $p = .018$) and *character* and *level of abstraction* ($F(1, 58) = 8.38$, $p = .005$). Significant main effects were obtained for *NPR style* ($F(3, 56) = 17.71$, $p < .0001$) and *level of abstraction* ($F(1, 58) = 425.19$, $p < .0001$). The resultant pattern corroborates and refines the findings from the MANOVA above, pointing out that recognizability and likeability were influenced by NPR style and level of abstraction (higher levels led to lower recognizability ratings), and both were differentially affected by the NPR style depending on the VC. Interestingly, membership of diagnostic group seems to have no influence on the ratings.

6 Conclusion

We investigated the differences between neuro-typically developed (NTD) children and adolescents, and those with ADHD and high-functioning ASD when categorizing stylized emotional facial expressions. Our first hypothesis was corroborated by the higher accuracy rates of the NTD subjects over those with ADHD and ASD. Our second hypothesis was that stylization and abstraction of visual emotion stimuli affects accuracy rates of ASD differently than ADHD and NTD. In this regard, we obtained valuable insights into the general perception of stylized emotion representations. Analyzing each group separately, the ASD group did not show significant preferences for any NPR styles over the realistic one. This could be understood as a lacking “abstracted faces effect”, which indicates the atypical processing of facial emotional stimuli in ASD. For the ADHD group, the medium abstracted CLD and IA styles obtained higher accuracy rates; however this was not consistent across all emotions and there was no style where ADHDs were as good or even outperformed NTDs, as formulated in our first hypothesis. Surprisingly, the reduction of details in the face due to the stylization did not improve accuracy for the ASD group.

An explanation for these results is provided by Zell et al. [2015], who found that blurring a realistic texture does not significantly reduce realism. This is demonstrated with the IA style, which contributed to higher accuracy rates, at least in the NTD group. Similarly, they observed that material stylization reduces the perceived intensity of expressions, explaining why emotions other than happiness were more difficult to categorize. Our findings regarding the low accuracy rates for categorization of fear, and to some extent for sadness and disgust, are consistent with previous works that suggested that these

Recognizability					
Style	VC	NTD	ADHD	ASD	Total
Ori.	M	2.6	2.23	2.46	2.41
	F	2.93	2.18	2.25	2.39
AI (M)	M	2.4	2.41	2.71	2.52
	F	2.47	2.41	2.33	2.39
AI (H)	M	2.67	2.82	2.75	2.75
	F	3	2.68	2.96	2.87
CLD (M)	M	3.4	4.09	3.63	3.74
	F	3.4	4.27	3.25	3.66
CLD (H)	M	3.53	4.86	4.29	4.31
	F	4	4.45	3.21	3.85
PD (M)	M	3.4	4	3.83	3.79
	F	4.33	4	4.17	4.15
PD (H)	M	3.67	4	3.96	3.9
	F	4.73	4.27	4.67	4.54
W (M)	M	3.87	3.68	3.79	3.77
	F	3.6	3.27	3.42	3.41
W (H)	M	4	4.45	4.33	4.3
	F	4.2	4.14	4.13	4.15

Likeability					
Style	VC	NTD	ADHD	ASD	Total
Ori.	M	2.87	2.41	2.63	2.61
	F	2.87	2.18	2.29	2.39
IA (M)	M	2.8	2.68	3.13	2.89
	F	2.87	2.64	2.67	2.7
IA (H)	M	3.4	3.05	2.79	3.03
	F	3.33	3.18	3.17	3.21
CLD (M)	M	4	3.95	4	3.98
	F	3.8	3.68	3.42	3.61
CLD (H)	M	3.8	4.45	4.42	4.28
	F	4.33	4	3.29	3.8
PD (M)	M	3.53	3.86	3.92	3.8
	F	4	3.45	4.33	3.93
PD (H)	M	3.67	3.91	3.71	3.77
	F	4.73	4	4.75	4.48
W (M)	M	3.93	3.86	4.17	4
	F	4.2	3.64	3.54	3.74
W (H)	M	4.2	4.55	4.63	4.49
	F	4.07	4.41	4.17	4.23

Table 3: *The participants rated recognizability (left) and likeability (right) of the rendering styles and virtual characters (VC) from 1 (very good) to 7 (very bad). Results are color-coded from white to red, more intense red meaning worse.*

emotions are the most challenging to recognize ([Harms et al. 2010], [Alves et al. 2013]). Pelphrey et al. [2002] also observed a deficit among autistic individuals in the identification of fear, and in general in tasks involving recognition of very simple or basic emotions. The recognizability-likeability questionnaire shed light on the perception of the NPR styles among all groups, showing that pencil based styles (CLD and PD) are harder to recognize and not as likeable as the realistic or more painterly (Watercolor) ones. One conclusion that might be drawn from this is that pencil based styles enhance, instead of reduce, the level of details in the face. The better recognition of emotions with the female character is also consistent with former research, where emotions were better recognized on female actors overall ([Battocchi et al. 2005], [Zibrek et al. 2013]). Another factor is the age difference between characters, with accentuated wrinkles and furrows in the elderly character when using pencil-based styles.

The obtained data is of great value as it delineates guidelines for our future work. First, we need to improve the NPR algorithms to achieve a more uniform level of abstraction that can be compared among the styles. Secondly, the animations should be artistically enhanced to accentuate the emotional meaning (e.g. add head or shoulder movement). We are also creating new characters that are the counterparts of the existing ones (an elderly female and a younger male) to better assess the influence of age and gender in both the NPR abstractions and emotion categorization. Finally, further studies need to be done to detect finer differences in the way the clinical groups perceive and categorize emotions. The fact that no significant difference was found among styles does not necessarily show that the participants were unaware of them. An hypothesis could be that the different stimuli (emotional expression, NPR style, age difference) blocked or dampened the impression of the facial expression, and thus the categorization of the emotions. This would be consistent with the view that NTD have a more holisitc processing of faces, while ASD tend to have fragmentary perception of faces, relying more on local facial elements than NTD [Deruelle et al. 2008]. In this sense, future experiments will refine the level of abstraction by selecting particular regions of the face (e.g. eye region, mouth region) based on eye-tracking information.

Acknowledgements

We would like to thank the German Research Foundation (DFG) for financial support within the projects AR 892/1-1, DE 620/18-1, and RA 764/4-1.

References

ALCORN, A., PAIN, H., RAJENDRAN, G., SMITH, T., LEMON, O., PORAYSKA-POMSTA, K., FOSTER, M. E., AVRAMIDES, K., FRAUENBERGER, C., AND BERNARDINI, S. 2011. Social communication between virtual characters and children with autism. In *15th Int. Conf. on Artificial Intelligence in Education*.

ALVES, S., MARQUES, A., QUEIRÓS, C., AND ORVALHO, V. 2013. Lifeisgame prototype: A serious game about emotions for children with autism spectrum disorders. *PsychNology Journal 11*, 3, 191–211.

AMERICAN PSYCHIATRIC ASSOCIATION. 2013. *Diagnostic and statistical manual of mental disorders: DSM-5 (5th ed.)*. Arlington, VA: American Psychiatric Publishing.

BARON-COHEN, S., GOLAN, O., AND ASHWIN, E. 2009. Can emotion recognition be taught to children with autism spectrum conditions? *Philosophical Transactions B*, 364, 3567–3574.

BATTOCCHI, A., PIANESI, F., AND GOREN-BAR, D. 2005. A first evaluation study of a database of kinetic facial expressions (dafex). In *Proceedings of ICMI 05*, ACM, 214–221.

BEHRMANN, M., THOMAS, C., AND HUMPHREYS, K. 2006. Seeing it differently: visual processing in autism. *Trends in Cognitive Sciences 10*, 6, 258–264.

BORA, E., AND PANTELIS, C. 2015. Meta-analysis of social cognition in attention-deficit/hyperactivity disorder (adhd): comparison with healthy controls and autistic spectrum disorder. *Psychological Medicine 46*, 4, 699–716.

BOUSSEAU, A., KAPLAN, M., THOLLOT, J., AND SILLION, F. 2006. Interactive watercolor rendering with temporal coherence and abstraction. In *Proceedings of NPAR 06*, ACM.

CARTER, E., HYDE, J., WILLIAMS, D., AND HODGINS, J. 2016. Investigating the influence of avatar facial characteristics on the social behaviors of children with autism. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.

COLLIN, L., BINDRA, J., RAJU, M., GILLBERG, C., AND MINNIS, H. 2013. Facial emotion recognition in child psychiatry: a systematic review. *Research in Developmental Disabilities 34*, 5, 1505–1520.

DERUELLE, C., RONDAN, C., SALLE-COLLEMICHE, X., BASTARD-ROSSET, D., AND DA FONSÉCA, D. 2008. Attention to low- and high-spatial frequencies in categorizing facial identities, emotions and gender in children with autism. *Brain and Cognition 66*, 2, 115–123.

EKMAN, P., FRIESEN, W., AND HAGER, J. 2002. *The Facial Action Coding System*. Weidenfeld & Nicolson, London, UK.

GEORGESCU, A. L., KUZMANOVIC, B., ROTH, D., BENTE, G., AND VOGELEY, K. 2014. The use of virtual characters to assess and train non-verbal communication in high-functioning autism. *Frontiers in Human Neuroscience 8*, 807.

GOOCH, B., AND GOOCH, A. 2001. *Non-Photorealistic Rendering*. A.K. Peters.

GRAWEMEYER, B., JOHNSON, H., BROSNAN, M., ASHWIN, E., AND BENTON, L. 2012. Developing an embodied pedagogical agent with and for young people with autism spectrum disorder. In *Intelligent Tutoring Systems*, Springer, LNCS, 262–267.

HAPPÉ, F., AND FRITH, U. 2006. The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *J Autism Dev Disord 36*, 1, 5–25.

HARMS, M. B., MARTIN, A., AND WALLACE, G. L. 2010. Facial emotion recognition in autism spectrum disorders: A review of behavioral and neuroimaging studies. *Neuropsychol Rev*, 20, 290–322.

HELZLE, V., BIEHN, C., SCHLÖMER, T., AND LINNER, F. 2004. Adaptable setup for performance driven facial animation. In *ACM SIGGRAPH 2004 Sketches*, 54.

HYDE, J., CARTER, E. J., KIESLER, S., AND HODGINS, J. K. 2013. Perceptual effects of damped and exaggerated facial motion in animated characters. In *FG*, IEEE Computer Society, 1–6.

KANG, H., LEE, S., AND CHUI, C. K. 2007. Coherent line drawing. In *Proceedings of NPAR 07*, ACM, 43–50.

KENNEDY, D., AND ADOLPHS, R. 2012. Perception of emotions from facial expressions in high-functioning adults with autism. *Neuropsychologia*.

KIRCHNER, J. C., HATRI, A., HEEKEREN, H. R., AND DZIOBEK, I. 2011. Autistic symptomatology, face processing abilities, and eye fixation patterns. *J Autism Dev Disord*, 41, 158–167.

KYPRIANIDIS, J. E., AND KANG, H. 2011. Image and video abstraction by coherence-enhancing filtering. *Computer Graphics Forum 30*, 2, 593–602.

KYPRIANIDIS, J. E., KANG, H., AND DÖLLNER, J. 2009. Image and video abstraction by anisotropic Kuwahara filtering. *Computer Graphics Forum 28*, 7, 1955–1963.

LOZIER, L. M., VANMETER, J. W., AND MARSH, A. 2014. Impairments in facial affect recognition associated with autism spectrum disorders: A meta-analysis. *Development and Psychopathology 26*, 4 Pt 1, 933–945.

LU, C., XU, L., AND JIA, J. 2012. Combining sketch and tone for pencil drawing production. In *Proceedings of NPAR 12*, 65–73.

LUFT, T., AND DEUSSEN, O. 2006. Real-time watercolor illustrations of plants using a blurred depth test. In *Proc. of NPAR 06*, 11–20.

MÄKÄRÄINEN, M., KÄTSYRI, J., FÖRGER, K., AND TAKALA, T. 2015. The funcanny valley: A study of positive emotional reactions to strangeness. In *Proceedings of the 19th International Academic Mindtrek Conference*, ACM, 175–181.

MCCLOUD, S. 1994. *Understanding Comics: The Invisible Art*. HarperPerennial.

MCDONNELL, R., BREIDTY, M., AND BÜLTHOFFYZ, H. H. 2012. Render me real? Investigating the effect of render style on the perception of animated virtual humans. *ACM Trans Graph 31*, 4.

MILNE, M., LUERSSEN, M., LEWIS, T., LEIBBRANDT, R., AND POWERS, D. 2011. 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*. IGI Global, 23–48.

MOTTRON, L., DAWSON, M., SOULIÈRES, I., HUBERT, B., AND BURACK, J. 2006. Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *J Autism Dev Disord 36*, 1, 27–43.

PAPARI, G., PETKOV, N., AND CAMPISI, P. 2007. Artistic edge and corner enhancing smoothing. *IEEE Trans Image Process 16*, 10, 2449–2462.

PELPHREY, K. A., SASSON, N. J., REZNICK, J. S., PAUL, G., GOLDMAN, B. D., AND PIVEN, J. 2002. Visual scanning of faces in autism. *J Autism Dev Disord 32*, 4, 249–261.

RAUH, R., AND SCHALLER, U. 2009. Categorical perception of emotional facial expressions in video clips with natural and artificial actors: A pilot study. Tech. rep., University of Freiburg.

ROSSET, D. B., RONDAN, C., DA FONSECA, D., SANTOS, A., ASSOULINE, B., AND DERUELLE, C. 2008. Typical emotion processing for cartoon but not for real faces in children with autistic spectrum disorders. *J Autism Dev Disord 38*, 5, 919–925.

SERRET, S., HUN, S., IAKIMOVA, G., LOZADA, J., ANASTASSOVA, M., SANTOS, A., VESPERINI, S., AND ASKENAZY, F. 2014. 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.

TARTARO, A., AND CASSELL, J. 2008. Playing with virtual peers: Bootstrapping contingent discourse in children with autism. In *Proceedings of the 8th International Conference on International Conference for the Learning Sciences*, ICLS'08, 382–389.

TATAROĞLU, E. 2013. A special experience in art education: Autism and painting. *EJBSS 1*, 11, 63–68.

ULJAREVIC, M., AND HAMILTON, A. 2013. Recognition of emotions in autism: a formal metaanalysis. *J Autism Dev Disord 43*, 7, 1517–1526.

VALLA, J. M., MAENDEL, J. W., GANZEL, B. L., BARSKY, A. R., AND BELMONTE, M. K. 2013. Autistic trait interactions underlie sex-dependent facial recognition abilities in the normal population. *Frontier in Psychology 4*, 286.

ZELL, E., ALIAGA, C., JARABO, A., ZIBREK, K., GUTIERREZ, D., MCDONNELL, R., AND BOTSCH, M. 2015. To stylize or not to stylize? The effect of shape and material stylization on the perception of computer-generated faces. *ACM Trans. Graph. 34*, 6, 184:1–184:12.

ZIBREK, K., HOYET, L., RUHLAND, K., AND MCDONNELL, R. 2013. Evaluating the effect of emotion on gender recognition in virtual humans. In *Proceedings of the ACM Symposium on Applied Perception*, ACM, SAP '13, 45–49.