SOCIAL PERCEPTION IN AUTISM
An eye tracking and pupillometric study

Master Degree Project in Cognitive Neuroscience
One year Advanced level 30 ECTS
Spring term 2018

Morgan Frost

Supervisors: Nouchine Hadjikhani
Jakob Åsberg Johnels
Martyna Galazka
Judith Annett

Examiner: Sakari Kallio

1Gillberg Neuropsychiatry Centre, University of Gothenburg, Gothenburg, Sweden
2University of Skövde, Skövde, Sweden
## Contents

1 Abstract 3

2 Background 4
   2.1 Social Processing 4
      2.1.1 Evolutionary Background 4
      2.1.2 The Social Brain Network 5
      2.1.3 Social Orienting and Attention 6
   2.2 The Pupil 10
      2.2.1 Pupil Anatomy 10
      2.2.2 The Pupil as a Marker of Attention Shifts 11
      2.2.3 Pupillometry 12
   2.3 Social Deficits in Autism Spectrum Disorder 13
      2.3.1 Social Orienting and Attention in ASD 13
      2.3.2 Neurophysiology of Social Processing in ASD 16
      2.3.3 Early Symptomatic Syndromes Eliciting Neurodevelopmental Clinical Examinations (ESSENCE) 21
   2.4 Aim of the Present Study 22

3 Methods 23
   3.1 Participants 23
   3.2 Clinical Metrics 24
   3.3 Apparatus 25
   3.4 Stimuli 25
   3.5 Procedure 27
   3.6 Data Analysis 27
   3.7 Ethics 28

4 Results 28
   4.1 Pupil Measures 28
   4.2 Eye Tracking Measures 29
      4.2.1 Latency 29
      4.2.2 Duration 31

5 Discussion 31

6 Conclusion 42

7 Acknowledgements 43

8 References 44
1 Abstract

Typically developing humans innately place subjective value on social information and orient attention to it. This can be shown through eye tracking and pupillometry, a method used to show attentional engagement. Social brain development and social preference is present from infancy, and is thought to rely on a carefully balanced network of neurotransmitters and neural connections. Autism Spectrum Disorder (ASD) presents altered neural systems which cause individuals to perceive and process social information differently, but the neurophysiology of this difference remains unclear. Previous research shows atypical gaze patterns, hyperarousal, and lack of orienting to social stimuli in ASD. Since autism is highly comorbid and shares traits with other neurodevelopmental disorders, it is difficult to distinguish aspects of these social processing differences. This study used a group of 35 neuropsychiatric patients to investigate how individuals with autism process social and non-social scenes. Eye tracking and pupillometry measures were collected while participants observed images of natural scenes with or without a person. Participants with autism did not show a pupillary response to social images and were slower to fixate on the face region than the other participants. Additionally there were correlations between clinical measures of social functioning and the length of time it took to fixate to faces. The results highlight important distinctions of social processing in autism. This thesis proposes a new perspective of looking at the social deficits present in autism spectrum disorder. It suggests reframing the current discussion from two leading hypotheses to a unified approach and formally considering the limitations of differing types of stimuli.

keywords: Autism Spectrum Disorder; ESSENCE; pupillometry; eye tracking; social processing; orienting; gaze aversion
2 Background

2.1 Social Processing

2.1.1 Evolutionary Background

Humans are a social species. We live and thrive in large social groups and exhibit collective behavior, which provides a number of evolutionary advantages. Social living yields learning opportunities for group members for strategies in how to use and cope with the environment. Living in large social groups also increases genetic diversity and protection from potential predators. The evolution of the human brain is very closely linked with the increasing complexity of social interactions and the cognitive abilities they require (Emery, 2000). Our brains evolved to be more social and supportive of different cognitive functions from those species which live in smaller groups. Examples of more complex social cognition include: processing of social hierarchy (storing and accessing information relating to others’ identity and position within the hierarchy); recognizing and mirroring others’ emotions (Theory of Mind and emotional contagion); communication through body language, facial expressions, and spoken language; and predicting others’ behavior, intentions, and trustworthiness. The primary cues for all of these processes rely on visual information, and in highly social animals such as primates, this reliance on visual information has led to an increasingly elaborate visual social processing system including more than 30 brain regions dedicated to visual processing (Adolphs, 2009; Emery, 2000). While there are a number of regions dedicated to other types of sensory social processing, this discussion will be confined to the networks and regions involved in processing visual social information.
2.1.2 The Social Brain Network

Somewhat recent findings pertaining to cognitive social development have led to the general agreement upon a network of subcortical and cortical structures involved in human social processing and interaction commonly referred to as the "social brain network" (Brothers, 1990). The social brain consists of a number of regions including but not limited to the fusiform gyrus, superior temporal sulcus, medial prefrontal and orbitofrontal cortex, and amygdala.

Cognitive processing within the social brain network is activated and modulated by perceived eye contact (Senju & Johnson, 2009). This “eye contact effect” occurs when perceived eye contact activates a “subcortical face detection pathway” including the superior colliculus, pulvinar, and amygdala (Senju & Johnson, 2009). This could be the reason why, when presented with a social stimulus regardless of context, typically developing humans tend to fixate first and longest on the eyes (Birmingham, Bischof, & Kingstone, 2009; Guillon, Hadjikhani, Baduel, & Roge, 2014; Senju & Johnson, 2009).

Neural activity in the various regions of the Social Brain reflect different aspects of social cognition; certain areas of the temporal lobe are involved in the processing of biological motion and others’ actions, and parts of the frontal cortex are implicated in representing the understanding of others’ goals and intentions (Adolphs, 2009). The amygdala is implicated in orienting, emotion recognition processes, and aggression and fear responses. Lesion studies showed that amygdala damage caused an impaired ability to recognize fear from facial expressions due to the inability to guide gaze and regulate visual attention to salient facial features such as the eyes (Adolphs, 1994). Furthermore, the same studies found the amygdala to be an important region in alerting and orienting, specifically in recognizing various stimuli as salient due to their associations with reward or punishment, whether or not they are unpredictable, and their sub-
jective importance (Adolphs, 1994). An additional study found that compared to predictable tones, unpredictable tones elicited greater amygdala activation (Adolphs, 2009). These studies show the amygdala plays an important role in social orienting and attention.

2.1.3 Social Orienting and Attention

Typically developing humans tend to focus their attention on the “T” shape on an individual’s face created by the eyes, nose, and mouth. The eyes reveal a lot of information about an individual’s emotional state, in addition to initiating processing within the social brain network. It is thought that the reason humans have a depigmented sclera (in comparison with other primates in which this does not exist) is due to the evolution of eyes for more effective social interactions and communication based on eye contact (Senju & Johnson, 2009). Furthermore, Åsberg Johnels et al. (2014) found that typically developing children pay attention to an individual’s mouth while he/she is talking, which could be important for language comprehension. Another group found that subjects pay attention to an individual’s nose when the individual is moving, suggesting the subjects used the nose as an anchor to follow movement (Vo, Smith, Mital, & Henderson, 2012).

Since the emergence of eye tracking technology, research has provided empirical evidence to show that humans have a basic tendency to attend to the central facial region, especially the eyes, of other humans when looking at faces (Birmingham, Bischof, & Kingstone, 2009; Pelphrey et al., 2002). Perhaps even more significantly, Birmingham, Bischof, & Kingstone (2009) found that this preference for face fixation was present even in complex social scenes, indicating that the bias for face gaze is, indeed, indicative of the central face as socially informative and communicative stimuli. This group also demonstrated the need for using scenes with more complex information instead of just faces in order to
ensure that subjects are indeed viewing central facial regions due to social bias and not due purely to salience (Birmingham, Bischof, & Kingstone, 2009).

When viewing more complex scenes, typically developing humans fixate first on any people within the scene, counting the number of people and reading their expressions and emotions. Then other objects in the scene are viewed with differing salience based on aspects such as contrast, color, intensity, size, and edge orientation (Henderson 2003). Eye tracking has also shown that humans orient and fixate first to the central face, prioritizing this social information over other information regardless of other, potentially salient, non-social information found within the scene (Birmingham, Bischof, & Kingstone, 2009). The amount of time it takes an individual to find their first fixation is known as latency. When presented with social images, humans tend to have very low latency to social stimuli (Crawford et al., 2016), which means they very quickly fixate on the social aspects of images.

The attentional bias for social stimuli is present early on in development. Recent studies have shown its presence as early as the third trimester, with fetuses showing preference for 3-dot infrared projections arranged in a pattern reminiscent of a face (two dots on top, one on the bottom), while showing no preference for other orientations (Reid et al., 2017). Indeed, already in 1968, it was shown that one-month-old infants also prefer social stimuli in the form of faces over non-social stimuli in the form of checkerboards, fixating first and longest on faces (Fitzgerald, 1968). This visual social preference has been noted in newborns and other infants for people, faces, and body motion as social stimuli (Chita-Tegmark, 2016; Gliga & Csibra, 2007). Infants have also shown pupillary contagion or pupil mimicry (Fawcett, Arslan, Falck-Ytter, Roevers, & Gredebäck, 2017), a preconscious phenomenon found in primates (Kret, Tomonga, & Matsuzawa, 2014) and humans in which pupils dilate in response to view-
ing another individual with dilated pupils potentially as a means of evaluating trustworthiness and emotion (Kret, Fischer, & De Dreu, 2015). The early development of basic social visual preference and subcortical processing indicates their ontogeny and evolutionary roots and implies adaptive functions and outcomes (Fawcett et al., 2017).

Social perception is a product of shifts between orienting and alerting and the subjective value innately placed on social stimuli which causes typically developing humans to actively shift attention toward social information (Bast, Poustka, & Freitag, 2018). Attention shifts are thought to arise due to activity of the subcortical structures locus coeruleus (LC) and amygdala and the neurotransmitter GABA.

The LC is located in the midbrain and acts as a “hub” for the noradrenergic system, modulating the transmission of norepinephrine (NE) to the rest of the brain (Bast, Poustka, & Freitag, 2018). The LC regulates sensory processing and attention through the release of NE, and consists of two distinct modes of activity, tonic and phasic. Phasic LC activity occurs when a subject is orienting to relevant stimuli or is actively focusing on a task. When the cells of the LC are in phasic mode, they activate in response to relevant stimuli. Phasic mode is also referred to as “exploitation mode” because the LC adjusts attentional filtering to ignore new stimuli to focus solely on the task at hand (Laeng, Sirios, & Gredebäck, 2012). Tonic LC activity, or “exploration” mode, occurs when the LC shifts the scope of attention between relevant events and does not maintain focus on one task. Subjects tend to exhibit poor task performance in tonic mode, but are more likely to detect novel aspects of the stimuli in tonic mode than in phasic mode (Laeng, Sirios, & Gredebäck, 2012). Social perception and cognition rely heavily on proper LC function which is based on an individual’s ability to not only appropriately shift attention between social and non-social
stimuli but also to maintain attention on relevant social stimuli (Bast, Poustka, & Freitag, 2018).

The amygdala, introduced above as a part of the social brain network, is a key component of social orienting. It is activated when a subject is shown social stimuli even when the stimuli are not consciously perceived (Herry, Bach, Esposito, DiSalle, & Perrig, 2007). The amygdala is thus implicated in rapid, automatic social processing such as bias and stereotypes, emotional face recognition, and evaluation of the salience of stimuli based on reward or threat (Adolphs, 2009). Indeed, amygdala lesion studies have shown atypical facial gaze patterns resulting in impaired fear recognition (Adolphs, 1994).

Another important attentional regulator is the neurotransmitter GABA. GABA is essential for proper neuronal proliferation, migration, differentiation, and other developmental processes. During early stages of neurodevelopment it acts as an excitatory neurotransmitter dependent on accumulation of chloride ions within neuronal cells (Cellot & Cherubini, 2014). As neuronal development progresses, the chloride is expelled from the cells and GABA becomes an inhibitory neurotransmitter (Le & Xu, 2008). One function of GABA as an inhibitory neurotransmitter is the regulation of pyramidal neuron activity in the prefrontal cortex (Markram et al., 2004; White, 1989). This function aids in the synchrony of gamma oscillations in cortical networks, including the networks involved in attentional processing (Bartos et al., 2007; Gruber et al., 1999; Steinmetz et al., 2000). Paine, Slipp, and Carlezon (2011) blocked GABA receptors in the prefrontal cortex of rats. The resulting hyperactivation caused attentional deficits, implicating proper GABA function as a component of attentional function. As an inhibitor, GABA also prevents hyperactivation in response to low-intensity stimuli (including social stimuli) to avoid unnecessary stress responses (Cellot & Cherubini, 2014).
2.2 The Pupil

2.2.1 Pupil Anatomy

Visual perception begins at the eye, within a pigmented structure called the iris. The iris contains two opposing groups of smooth muscle, called the sphincter and dilator, as well as an opening called the pupil, through which light enters the eye. The diameter of the pupil changes rapidly, and can vary from less than 1 mm to more than 9 mm (Beatty & Wagoner, 2000). Pupil constriction and dilation is regulated by the antagonistic parasympathetically-controlled sphincter and sympathetically-controlled dilator muscles, respectively, which surround it. Activation of the dilator muscles due to increased sympathetic activity inhibits parasympathetic activity by reducing motor output to the sphincter muscles in a complex, reciprocal system initiating from the parasympathetic motor nucleus in the midbrain. The size of the pupil at any given moment is a result of a ratio of activity occurring in the neural pathways which control these muscles (Beatty & Wagoner, 2000).

Pupillary movements can serve primarily optical purposes, such as the pupillary light reflex. This reflex is a regulation of the amount of light entering the eye, proportional to the diameter of the pupil which is constricted or dilated in bright light or dim light, respectively. In addition, there are pupillary reflexes which accommodate to visual stimuli that are unfocused or binocular due to nearness in proximity, as well as reflexive dilation which is evoked in response to sensory occurrences. However, pupillary constriction and dilation do not only occur in the presence of external sensory stimuli. Indeed, there are very small (less than 0.5 mm) changes in pupillary diameter that are related to cognition and emotion processing (Beatty & Wagoner, 2000). These fluctuations appear to serve no optical function. Rather, they reflect changes in brain activation that underlie cognitive events (Bradley, Miccoli, Escrig, & Lang, 2008).
For this reason they can serve as empirical indicators and in certain cases even biomarkers of cognitive brain processes and thus are of great interest to the field of cognitive neuroscience.

### 2.2.2 The Pupil as a Marker of Attention Shifts

The LC is thought to be an essential area within the neural system which also controls the muscles in the iris (Laeng, Sirios, & Gredebäck, 2012). This relationship provides a way for researchers to observe activational changes within the LC through pupil dilation, an externally measurable response to internal LC activation. Thus, this LC-NE-pupil connection means researchers can measure pupil diameter in order to observe shifts in attention. This is supported by decades of studies which have shown that pupil diameter is directly correlated with activational changes in LC neurons (Laeng, Sirios, & Gredebäck, 2012; Rajkowski, Kubiak, & Aston-Jones, 1993; Rajkowski, Majczynski, Clayton & Aston-Jones, 2004). Even before the LC-pupil connection was established, Sokolov (1963) indicated that pupil dilation was a component of attention orienting to novel stimuli.

Indeed, studies have established that changes in pupil diameter predict changes in tonic and phasic LC activity: during low tonic state, the pupil fluctuates while slowly decreasing in diameter, while during phasic LC activity there are large pupillary increases (Laeng, Sirios, & Gredebäck, 2012). Particularly salient stimuli also cause pupil dilation (Laeng, Sirios, & Gredebäck, 2012). Rajkowski, Kubiak, and Ashton-Jones, (1993) also showed the link between the LC-NE system and pupillary responses in non-human primates. Additionally, Einhauser, Stout, Koch, & Carter, (2008) used bistable images (two different percepts within one image, such as the Necker Cube) to observe pupillary changes alongside shifts in perception. Asking subjects to press a button at the time of perceptual switching revealed that pupil diameter increased significantly
around 100 ms before the perceptual shift and the larger the pupil diameter at this time, the more stable the perception subsequently (Einhauser et al., 2008). This demonstrates the role of pupillary responses in attentional shifting and recognition of relevant events.

2.2.3 Pupillometry

Given that social processing is a product of LC-NE activity and attentional shifts, pupil diameter can therefore be useful for observing social processing (Bast, Poustka, & Freitag, 2018; Fitzgerald, 1968). The measurement of pupil diameter to determine underlying cognitive events is called pupillometry.

Pupillometry has been used in clinical settings for several decades for psychiatric, neurological, and medical functions. This includes diagnostic purposes, monitoring anesthetic effects, assessing autonomic function, and more recently to determine cognitive load (Graur & Siegle, 2013).

Pupillometry is a particularly useful tool in cognitive neuroscience research because it is a measure of preconscious processing and cognitive events, while in addition it is non-invasive and relatively inexpensive. It is especially useful in nonverbal populations such as infants or severely developmentally disabled patients, since it does not require conscious, verbal response. Finally, it is also especially useful when studying patients with autism given that it is minimally stimulating and thus is not likely to cause anxiety and the resulting potentially interfering brain signals (as opposed to loud, confined machines as in functional magnetic resonance imaging (fMRI), or novel physical sensations such as electroencephalography (EEG)).
2.3 Social Deficits in Autism Spectrum Disorder

Autism Spectrum Disorder (hereafter abbreviated as ASD) is a neurodevelopmental disorder characterized by a number of social and communication deficits, restricted interests, and repetitive behaviors as recognized in the DSM-V (American Psychiatric Association, 2013). Among the most prominent social deficits in ASD is the reduced ability to use social cues from facial expressions and body language for learning and social interactions. Such deficits may stem from basic attentional processing deficits present in ASD (Bast, Poustka, & Freitag, 2018) or from an excitation and inhibition imbalance caused by abnormalities in GABA function (Hadjikhani et al., 2017).

2.3.1 Social Orienting and Attention in ASD

Atypical facial gaze patterns and eye contact are among the most discernible indications of social impairments in ASD (Pelphrey et al., 2002; Senju & Johnson, 2009). This phenomenon can be shown through the use of eye tracking (Boraston & Blakemore, 2007). It has been observed in patients with ASD since the 1940s (Kanner, 1944) and is included in standardized diagnostic criteria including the DSM-V (American Psychiatric Association, 2013). Atypical face scanning in ASD patients can be observed as early as infancy (Hutt & Ounsted, 1966), even prior to diagnosis in high-risk individuals (Senju & Johnson, 2009), and is present throughout their lifespan. As seen in Figure 1, in comparison with controls who scan a strategic, predictable triangle composed of the eyes, nose, and mouth, the scanpath of ASD subjects faces tends to be “erratic, undirected, and disorganized” (Pelphrey et al., 2002). Nearly all studies using isolated face stimuli observe this phenomenon.

When using more complex social scenes as stimuli, however, results have been inconsistent. The inconsistency in eye tracking results might indicate a
Figure 1: Scanpath examples of subjects with ASD and typically developed subjects (from Pelphrey et al., 2002).

larger issue in autism research, that the specific characteristics of the stimuli presented could have a large impact on the outcome of the study. Some studies find that individuals with ASD attend far less to social stimuli even with more ecologically valid stimuli. For example, Crawford et al. (2016) found that, when shown two simultaneous videos of geometric shapes and social scenes, toddlers with ASD fixated significantly more upon the geometric shapes, whereas both typically developing and developmentally delayed toddlers preferred the social videos. Furthermore, toddlers with ASD did not show preferential attention for human biological motion in comparison with both typically developing and developmentally delayed control groups (Crawford et al., 2016). Riby and Hancock (2008) found that, in scenes containing people, autistic subjects with an average
mental age of 5 years fixated to the eye regions less than typically developing subjects of the same age.

However, there are also many studies which have found the opposite – that individuals with ASD show no difference from TD individuals in gaze patterns of complex social scenes. For example, Chawarska et al. (2012) found that when viewing a movie of an actress making a sandwich, typically developing toddlers and those with ASD spent the same amount of time looking at the actress’ face. Freeth et al. (2010) also found a similar looking time to faces in TD adolescents and adolescents with ASD.

Klin et al. (2002) used film clips of “intense social interactions” to show that ASD subjects fixated to the eye regions for significantly less time than controls. However, Speer and colleagues (2007) used still frames from the same clips as Klin et al. (2002), and found that high-functioning autistic subjects showed the same gaze duration as typically developing controls when audio, interaction, and motion information were removed, but differed significantly when the dynamic videos were maintained.

Toddlers with ASD have also shown fewer attentional shifts from face-to-face and object-to-face than typically developing and developmentally delayed controls (Sweetenham et al., 1998). Additional social attention deficits in ASD include lack of social smiles, impairments in joint attention, and decreased orienting to familiar social auditory stimuli, especially vocalizations of the individual’s name (Anderson, Colombo, & Shaddy, 2006; Dawson et al., 1998).

Compared to eye tracking, there have been relatively few studies which used pupillometry to observe social processing in autism. Anderson, Colombo, & Shaddy (2006) examined pupillary responses to face and non-face stimuli in autistic and typically developing groups. Whereas typically developing controls exhibited significantly larger pupil diameter in response to faces, the autistic...
group showed a significant decrease in pupil size in response to faces. In addition, the autistic group showed significant pupillary constriction in response to internal facial features (eyes, nose, and mouth) as opposed to external features (such as the chin or ears), in comparison with typically developing controls who displayed pupil dilation to both internal and external facial features (Anderson, Colombo, & Shaddy, 2006). In comparison to typically developing controls, adolescents with autism showed decreased pupil diameter when watching dynamic social videos (Muller, Baumeister, Dziobek, Banaschewski, & Poustka, 2016). Falck-Ytter (2008) demonstrated that 5-year-olds with ASD showed increased phasic pupil dilation in response to inverted faces compared to upright faces, whereas TD controls showed the opposite response. Children with ASD also did not show increased phasic pupil dilation in response to objects with emotional salience in contrast to TD controls (Nuske, Vivanti, & Dissanayake, 2016). More generally, Anderson and Colombo (2009) reported increased tonic pupil dilation in 2-6-year-olds with ASD, and Nystrom and colleagues (2015) found that 10-month-old infants at high risk for ASD showed a stronger and faster pupillary light reflex than low-risk infants, indicating a general NE abnormality and heightened autonomic activity in individuals with ASD. Hence, although pupillometry is still a not a widely used methodology in ASD research, previous research clearly indicate that such data can shed new light on many basic aspects of social perception.

2.3.2 Neurophysiology of Social Processing in ASD

Impairments as well as atypical activation and anatomy within the social brain network are apparent in individuals with ASD (Hadjikhani, Joseph, Snyder, & Tager-Flusberg, 2007; Johnson et al., 2005); however, the causes and mechanisms of impaired social cognition in ASD are not well understood. Currently there are two competing, and somewhat controversial explanations, both of
which have empirical support from studies using different methods and authors spanning over a decade, and neither of which is completely understood. These are commonly described as the “social motivation” hypothesis or “gaze indifference”, and the “gaze aversion” hypothesis. A fundamental challenge in autism research is to disentangle these two hypotheses empirically and to evaluate the proper methods with which to do so.

The “social motivation hypothesis” maintains that individuals with ASD are emotionally indifferent to social stimuli, stemming from social impairments which lead to a lack of information perceived from social cues (Chevalier, Kohls, Troiani, Brodkin, & Schultz, 2012). Typically developing people attend to social stimuli from a very early age, and from the perceived social information it is learned that social interactions are intrinsically rewarding. This process likely causes humans to modulate behavior in response to maintain and improve interpersonal relationships. Impaired social attention from an early stage of development could cause patients with ASD to miss necessary social cues which form the foundation of social cognition, cascading to cause these individuals to obtain little social reward from salient social stimuli and thus fail to learn from social interactions and cues (Bast, Poustka, & Freitag, 2018; Chevalier et al., 2012).

Using fMRI, Pelphrey et al. (2005) measured activation in the superior temporal sulcus (STS), a region most implicated in gaze processing, in response to gaze shifts in people with ASD and typically developing controls. They found that the ASD group did not differ in STS activation between a direct gaze and an averted gaze, indicating a diminished ability to perceive gaze as intentional (Pelphrey et al., 2005). There was also no difference in fear processing activation when patients with ASD were shown images of fearful faces with direct and averted gazes (Davies, Dapretto, Sigman, Sepata, & Bookheimer, 2011),
suggesting a lack of emotional relevance processing. Additionally, when directly cued to look at the eyes, a small group of two-year-old children with ASD did not look away faster than typically developing children, suggesting a lack of attention to the eye region rather than an intentional aversion (Moriuchi, Klin, & Jones, 2017).

Impaired LC function is a potential underlying mechanism which could help to maintain the social motivation hypothesis. Atypical pupil dilation in response to salient stimuli provides empirical evidence which implicates the LC in aberrant social attention in individuals with ASD (Anderson, Colombo, & Shaddy, 2006; Bast, Poustka, & Freitag, 2018; Muller et al., 2016 Sepata et al., 2012). Additionally, fMRI images of patients with ASD reveal atypical activation of functional brain clusters modulated by the LC-NE system (Bast, Poustka, & Freitag, 2018). However, as will be discussed below, the decreases in activation and pupil diameter could merely be a reflection of the avoidance of the central facial regions, since none of the studies controlled for gaze fixation. Lack of attention to social stimuli could be due to a learned avoidance of social stimuli, which could still be reflected in the LC-NE system.

Typically developing humans show reward anticipation and processing when presented with social stimuli (Sepata et al., 2012). Dawson, Bernier, & Ring (2012) suggest that impaired reward processing could also play a role in social deficits in ASD, with some research indicating a reduced reward sensitivity to social stimuli in individuals with ASD compared to TD individuals. The neural mechanisms of this impaired reward processing have been observed with fMRI and ERP studies. As measured with fMRI, the anterior cingulate cortex, implicated consistently in processing the reward values in drugs, food, sex, and indeed in the viewing of social stimuli, has shown reduced activation in individuals with ASD in response to social stimuli but not in response to other
intrinsically rewarding stimuli (Dawson, Bernier, & Ring, 2012). In an incentivized go/no-go task, children with ASD showed an attenuated P3 amplitude response to both social and non-social stimuli, indicating a deficit in general reward processing, not specific to social information (Dawson, Bernier, & Ring, 2012). Stavropoulos & Carver (2014) measured the stimulus preceding negativity (SPN) as implicated in reward anticipation, and the feedback related negativity (FRN) as implicated in reward processing. They found that children with ASD had smaller responses in both SPN and FRN accompanied by social stimuli compared to typically developing controls; however, when accompanied by non-social stimuli, there was no difference in either ERP between ASD and typically developing groups (Stavropoulos & Carver, 2014). Impaired social reward processing could explain a lack of orienting to social information in ASD; if individuals with ASD do not feel rewarded when attending to social information, they would orient less frequently to social stimuli (Chevalier et al., 2012). However, lack of reward processing could also be a result of hyperactivation, since a stimulus causing overarousal would be seen as unfavorable and thus not elicit a reward processing response.

The “gaze aversion hypothesis” infers that many differences in social processing, including face and gaze fixation and eye contact in individuals with ASD could be due to an aversion to looking at faces due to a heightened autonomic response to social stimuli, causing a fear response in ASD individuals which leads them to actively look away from the eyes (Dalton et al., 2005).

In line with such an over-arousal account, ASD patients show hypersensitivity and poor emotion regulation to low intensity fear faces (Lassalle et al., 2017). Moreover, when controlled for gaze fixation, looking at the eyes caused increased amygdala activation (Dalton et al., 2005) as well as pulvinar and superior colliculus (Hadjikhani et al., 2017) in patients with ASD, indicating
hyperactivation and possibly even a fearful response to eye contact. Additionally, subjects with ASD showed larger skin conductance response (SCR) than controls in response to social stimuli, indicating autonomic arousal (i.e. an emotionally aversive response) (Joseph, Ehrman, McNally, & Keehn, 2008). From this perspective social disengagement in ASD can be viewed as a compensatory (or protective) behavior to reduce feelings over over-arousal (Hadjikhani et al., 2017).

Porges’ polyvagal theory (2007), a subcategory of hyperarousal models of social processing, suggests that social processing difficulties in ASD are the result of an atypical regulation of the autonomic nervous system (ANS). Porges indicates that, in typical development, the parasympathetic and sympathetic branches of the ANS promote appropriate physiological responses by modulating one another in a balance. When parasympathetic control is increased, sympathetic arousal is decreased, causing a calmer physiological state and thus more controlled and appropriate social behavior (Porges, 2007). In ASD, an imbalance in these systems would cause a tense ANS and improperly controlled social responses. Additionally, sympathetic and parasympathetic balance is implicated in pupil dilation and constriction and could possibly be shown through pupillometry.

Gaze aversion could also be neurophysiologically explained by improper GABA function in individuals with ASD. A leading hypothesis infers that some ASD symptoms arise from an imbalance between excitation and inhibition in neural circuits involved in sensory, social, and emotional processes (Pizzarelli & Cherubini, 2011). The cortex is organized in columns of sensory processing neurons, controlled by GABA through lateral inhibition. Analysis of postmortem tissues from ASD patients has revealed alterations in the number, spacing, and structure of these columns, which is associated with altered GABAergic sig-
naling (Pizzarelli & Cherubini, 2011). ASD animal models (created by alterations of known autism candidate genes) have shown these structural alterations in neuronal columns in addition to dysfunctional GABAergic signaling, which caused hyper-excitability within the hippocampal network (Pizzarelli & Cherubini, 2011). In a recently published clinical trial, Hadjikhani et al. (2018) showed that with use of the diuretic Bumetanide, which decreases intracellular chloride and increases inhibition by GABA, ASD patients increased the amount of time they spent looking in the eye regions of social images which previously evoked hyperactivity in the amygdala, and their amygdala activation normalized. This is further evidence that improper GABA function, and the resulting excitation/inhibition imbalance, is a reputable potential explanation for the hypersensitivity to social stimuli experienced by individuals with ASD. It also provides a promising direction for future research.

2.3.3 Early Symptomatic Syndromes Eliciting Neurodevelopmental Clinical Examinations (ESSENCE)

Early Symptomatic Syndromes Eliciting Neurodevelopmental Clinical Examinations (ESSENCE) is a classification which refers to a group of neurodevelopmental disorders that are often comorbid, share symptoms, and are difficult to differentiate early in development. Some of these disorders are ASD, attention deficit hyperactivity disorder (ADHD), Tourette Syndrome, bipolar disorder, epilepsy, developmental coordination disorder (DCD), dyslexia, and several other early-onset neurological disorders presenting cognitive problems (Gillberg, 2010). These disorders are difficult to differentiate during early diagnosis due to shared symptoms, genes, and environmental risk factors. The disorders are also comorbid, meaning a child often has more than one disorder, and the symptoms can overlap with one another. Due to these issues, a child with ESSENCE is often misdiagnosed early on or only diagnosed with (and helped with) some of
the presenting problems (Gillberg, 2010). ESSENCE is not a specific diagnosis, rather it is an umbrella term highlighting the need for awareness of the variety in problems and needs of patients with ESSENCE. Since symptoms are shared across different disorders, it is often difficult to assess which behaviors and neurological factors arise from the different disorders. Social processing is one area of development which is affected by different disorders under the ESSENCE umbrella; thus, since patients with an ASD diagnosis often have comorbid diagnoses, it is difficult but essential to differentiate the neurological underpinnings of social processing in ASD from that of other neurological disorders.

2.4 Aim of the Present Study

This current study examined the visual scanning and pupil dilation in response to social and non-social images of a group of adolescents with either a diagnosis of ASD or another ESSENCE diagnosis. A clinical control group was used instead of a typically developing control group in order to outline specific differences in ASD, giving a clearer picture of ASD for diagnostic and research purposes.

Given the existing current literature indicating the extent of social impairments in ASD, it was expected that the ASD group would show differences in scene scanning from those without ASD. Specifically, it was hypothesized that ASD patients would take longer to fixate to, and spend less time fixating on, the social regions of the images, especially the central face. It was also hypothesized that the ASD patients would differ from participants without ASD diagnosis in pupil dilation when shown social images; specifically, that they would show less pupil dilation or even pupil constriction in response to social stimuli than their non-ASD counterparts. Finally, a correlational investigation was carried across diagnoses to examine the extent to which a patient’s social symptom severity
and functioning scores would associate with their scanning and pupil deficits.

3 Methods

3.1 Participants

The study group consists of 37 patients who had been referred to the Child Neuropsychiatric Clinic (CNC) in Gothenburg, Sweden between 2005 and 2009 and had been diagnosed with either ASD or another ESSENCE disorder(s) as children at the time of initial diagnosis. All diagnoses were based on criteria from the DSM-IV. A follow-up evaluation was carried out at the time of this study in which patients were re-evaluated for a variety of factors including diagnosis, IQ, social functioning, cognitive functioning, and well-being.

One participant was excluded from the study due to missing diagnostic information, and another was excluded due to insufficient measurement of pupil diameter, leaving the final sample size for this study at $N = 35$.

The age range for the 35 participants in the study group was 12 – 20 years ($m = 16.45, sd = 1.65$). The IQ range for this group, based on either the Wechsler Intelligence Scale for Children - Fourth Edition (WISC-IV) (Wechsler, 2003a) or the Wechsler Adult Intelligence Scale (WAIS-IV) (Wechsler, 2003b), was 65 – 120 ($m = 90.12, sd = 12.69$). The subjects were categorized according to two main diagnostic groups: ASD ($N = 22$) and non-ASD ($N = 13$). No difference was found between the two groups regarding age or IQ. Seven of the subjects (20%) were female, and 28 (80%) were male; the male:female ratio was 5.75:1 for ASD and 4.33:1 for non-ASD. In clinically referred children with ESSENCE the male:female ratio is typically 7-10:1 (Davidsson et al., 2017).
3.2 Clinical Metrics

The Five to Fifteen (FTF/5-15) is a questionnaire developed to assess development and behavioral difficulties in children aged 5 – 15 years (Kadesjö et al., 2004). Questions on the FTF are given to parents and fall into eight domains: motor function, attention/executive function, language, memory, learning, social competence, internalizing and externalizing behavior problems. For the purposes of this study, only the social competence scores were used in data analysis. Items are scored from 0 (does not apply), 1 (applies sometimes or to some extent), or 2 (definitely applies); and thus, the higher the number of a patient’s score, the lower the functioning of the child within the given domain. In this study sample, FTF social competence scores in the ASD group ranged from 0.19 – 1.37 ($m = 0.828; sd = 0.343$) and in the non-ASD group scores were 0.04 – 0.63 ($m = 0.214; sd = 0.212$). The difference was significant ($p < .001$) between groups, which is to be expected since a key diagnostic factor in ASD is poor social functioning.

The Global Assessment of Functioning (GAF) is a clinical assessment of a patient’s overall functioning level, on a scale ranging from 0 (inadequate information) to 100 (superior functioning) (Yamauchi, Ono, Y & Ikegami, 2001). Physical and environmental limitations are not assessed; rather it is a scale of a patient’s social, communication, psychological, and occupational functioning. In the ASD group, the GAF scores ranged from 45 – 65 ($m = 54.83; sd = 6$). In the non-ASD group, GAF scores ranged from 45 – 75 ($m = 62.27; sd = 10$). There was a between-group difference ($p < .05$) for GAF scores in this sample, with the ASD group overall having lower functioning than the non-ASD group.

The Vineland Adaptive Behavioral Scale (VABS) is a ranking of adaptive functioning given by a patient’s caregiver, and covers four general domains: Daily Living, Motor Functioning, Communication, and Socialization. This
study utilized the patients’ social scores on the VABS scale (Sparrow, Cicchetti, & Balla, 2005). In the ASD group, the VABS social scores ranged from 56 – 107 \( (m = 80.48; sd= 15.52) \). The non-ASD group scores ranged from 92 – 113 \( (m = 107.25; sd = 7.98) \). There was a between-group difference on this scale as well \( (p < .001) \), again expected due to the overall low social functioning that is a defining characteristic of ASD.

3.3 Apparatus

Gaze and pupil information were measured with a Tobii T120 (Tobii Technology Inc., Stockholm, Sweden), which records near infrared reflections of both eyes at 60 Hz as the subject watches an integrated 17 inch monitor. Subjects were seated 60 cm away from the monitor. A 9-point calibration procedure was used before the experiment, in which a small animated object is shown to the participant and collects data at nine points on the screen, repeating until data is collected for all points to ensure proper calibration.

3.4 Stimuli

There were a total of 22 images used in the study, divided into 11 pairs. Each pair consisted of the same scene with (social) and without (non-social) a person present in the image. Other than presence or absence of a person, the scenes within each pair were identical (see figures 2 and 3). Subjects from each diagnostic group were randomly divided into two groups, version A and version B, and shown a series of 11 images. Version A contained six social images and five non-social images, and version B contained five social images and six non-social images. No subject was shown two images within the same pair to avoid conflicting gaze patterns due to the subject recognizing the image. Each image was shown for 12 seconds.
Figure 2: An example of the social version of images used in this study. The social area of interest occupies the 3-5% of the image around the face.

Figure 3: An example of the non-social version of images used in the study. The image is exactly the same as its social counterpart (figure 2), with the only difference being the absence of a person.

To obtain the fixation measures, each image pair was divided into two areas of interest (AOIs) with Tobii studio: social and non-social. Social AOIs were created on the images that contained people and consisted of the face, composing 3 – 5% of the total image. non-social AOIs were created on the non-social images, and consisted of the regions (3 – 5% of the image) which corresponded to the social AOIs on the social version of the same image.
3.5 Procedure

After arrival at the Gillberg Centre, which was familiar to the subjects due to prior visits, subjects sat 60 cm in front of a monitor in the testing room. The equipment was adjusted to obtain valid pupil and eye tracking measurements. Subjects were instructed to look at pictures on the monitor freely, and the stimuli were presented for 12 seconds each.

3.6 Data Analysis

Pupil data were exported from Tobii studio and reduced using TimeStudio extension for MATLAB (Nyström, Falck-Ytter, & Gredebäck, 2016), which is an open-access analysis tool designed specifically for analyzing timeseries data. Data were inspected for various artifacts such as blinks and head movements, and rejected if these artifacts were present for more than 50% of the trace duration within the analyzed time frame (0 – 5 s). Pupil size was averaged for each subject across both eyes. To eliminate individual differences within subjects and trials, the data were adjusted and normalized according to a baseline calculation over the initial first second (0 – 1 s) after stimulus onset. Mean pupil size for each subject was then calculated over social and non-social images for the period of 1 – 5 seconds. These data were imported into SPSS for analysis. A related samples Wilcoxon Signed Rank test was performed within each group to test a difference between social and non-social pupillary response, and a Mann Whitney U test was performed to test for between-group differences. Spearman’s rank correlations were used to test the relationships between pupil data and eye tracking data as well as clinical metrics.

Eye tracking data were exported from Tobii studio for latency and total fixation duration within the social and non-social AOIs and analyzed using SPSS software. Mann Whitney U tests were run to test for between-group differ-
ences. Wilcoxon Signed Rank tests were run to test for within-group differences. Spearman’s rank correlations were used to test the relationships between latency and/or fixation duration and the clinical metrics.

Nonparametric tests were used throughout analysis to mitigate the effect of the small sample size and any outliers.

3.7 Ethics

Parents provided written consent. The study was approved by the Ethics Committee at the University of Gothenburg and performed in accordance with ethical standards of the institutional research committee. No conflicts of interest have been declared by any of the researchers associated with the study.

4 Results

4.1 Pupil Measures

A Wilcoxon signed-rank test was run to determine if there was a within-group difference in each group for pupil size in response to social and non-social stimuli. In the ASD group (mean social pupil = 0.029 mm, $sd = 0.156$; mean non-social pupil = -0.007 mm, $sd = 0.119$) the test showed no significant difference ($p = 0.394$). In the non-ASD group (mean social pupil = 0.084, $sd = 0.145$; mean non-social pupil = -0.042, $sd = 0.089$) a difference was found ($p = 0.019$) such that there was larger mean pupil size for social than non-social scenes. (see fig. 4 and 5).

A Mann-Whitney U Test was run to test for between-group differences in pupil size. No differences were found between social or non-social pupil size across groups (social pupil $p = 0.834$; non-social pupil $p = 0.139$). (see fig. 6)
4.2 Eye Tracking Measures

4.2.1 Latency

A Mann-Whitney U Test was run to determine whether the latency of the first fixation within the social AOI (social latency) was the same between both groups (ASD mean = 0.873 s, sd = 0.799; non-ASD mean = 0.453 s, sd = 0.201). A significant difference was found ($p = 0.043$) indicating that the group with ASD were slower than the non-ASD group at fixating to the social aspect of the stimulus.

Spearman’s correlations were run to assess the relationship between social latency and various clinical measures (see figures 7 – 9). Correlations were found between the latency to social stimuli and the clinical measures of GAF ($r = -0.5$, $p = 0.004$), VABS social ($r = -0.552$, $p = 0.003$), and social competence scores


\[(r = 0.446, \ p = 0.012)\] for all participants.

Figure 5: Linear relationship between social latency \((s)\) and GAF, including best fit line and 95\% confidence interval.

Figure 6: Linear relationship between social latency \((s)\) and VABS social, including best fit line and 95\% confidence interval.
4.2.2 Duration

A Mann-Whitney U Test was run to determine differences for mean total duration of fixations for social AOIs (faces) between the ASD group (mean total duration = 0.193 s, $sd = 0.153$) and the non-ASD group (mean total duration = 0.284 s, $sd = 0.170$). No difference was found, but there was a strong trend ($p = 0.075$).

A Spearman’s correlation showed a positive linear relationship between total duration (s) to faces and difference in pupil diameter (mm) between social and non-social images (see fig. 8). ($r = 0.337, p = 0.028$).

5 Discussion

This study aimed to investigate the visual scanning and pupil dilation in response to social and non-social images in a clinical group of adolescents with ESSENCE, divided into groups based on presence or absence of an ASD diag-
Figure 8: Linear relationship between total duration (s) to faces and difference in pupil diameter (mm) between social and non-social images, including regression line and 95% confidence interval.

Based on findings from previous research, it was expected that the ASD group would respond less to social stimuli but their arousal-related reactions to social stimuli were hard to predict from previous research, which have provided evidence in favour of both reduced social motivation and in favour of theories of over-sensitivity to social information in ASD. The study also made comparisons between clinical measures and empirical measures to determine whether patients’ symptom severity correlated with their response to social stimuli.

The results presented here show that while the non-ASD group on average showed an increased pupil diameter in response to social stimuli and decreased pupil diameter in response to non-social stimuli, the ASD patients showed no difference in pupil response between social and non-social stimuli. In addition, the ASD group showed a longer latency to social information than the non-ASD group but did not differ in the amount of time spent fixating to social information. These results indicate a slower orienting response to social information by the ASD group and present possible markers for autism among other clinical diagnoses. The findings also suggest a correlation across diagnoses between social
clinical measures and latency to social orientation, providing a potential early marker for ESSENCE and a direction for future clinical research. See figure 9 for a summary of these findings and their outcomes.

<table>
<thead>
<tr>
<th>Findings</th>
<th>Meanings</th>
<th>Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>The ASD group showed no difference in pupil diameter between social and non-social stimuli; the non-ASD group had a larger pupil size in response to social images.</td>
<td>The ASD group showed reduced attention to social scenes and a lack of recruitment of the LC-NE system.</td>
<td>Atypical social orienting and attention in ASD. Papillometry as a potential sensitive and early marker for ASD.</td>
</tr>
<tr>
<td>The ASD group took twice as long to fixate to social AOIs as the non-ASD group.</td>
<td>Indication of attention/orienting dysfunction in the ASD group, as a possible outcome of atypical social brain development.</td>
<td>Latency as a distinguishing factor of social processing in ASD compared to other ESSENCE diagnoses.</td>
</tr>
<tr>
<td>Similar durations of social fixations across groups.</td>
<td>Overall, the two groups (ASD and non-ASD) spent similar amounts of time looking at the social AOIs.</td>
<td>The data on social fixation duration could not be used to differentiate between ASD and other ESSENCE conditions in this study.</td>
</tr>
<tr>
<td>Correlation between duration to social AOIs and difference in pupil diameter (social - non-social) across groups.</td>
<td>The more time a participant spent looking at social AOIs, the greater their pupil dilation — orienting &amp; attention — in response to social stimuli.</td>
<td>The type of stimulus used in the experiment evokes a different response from other types of stimuli used in various other experiments.</td>
</tr>
<tr>
<td>Correlations between clinical scores and social latency across groups.</td>
<td>The less socially functioning the patient (as evaluated by various clinical evaluations), the longer it took him or her to fixate to social AOIs.</td>
<td>Reinforces the validity of these clinical evaluations. Potential early marker for ESSENCE and social functioning, allowing for earlier intervention. Potential indicator of a common neurophysiological factor related to social orienting and attention in ESSENCE.</td>
</tr>
</tbody>
</table>

Figure 9: Table showing a summary of the findings presented here and their meanings and implications.

For the ASD group, there was no significant difference in pupil size between social and non-social stimuli, indicating reduced attention and a lack of recruitment of the LC-NE system when viewing these stimuli. Since pupil size differences were present in the non-ASD group, the pupillary response could be a potential sensitive indicator for autism. Furthermore, as pupillary responses to social stimuli are present early in infancy, they could serve as an early marker of ASD.

The ASD group took twice as long to fixate to social information within the images as the non-ASD group. The slow fixation is a further indicator of attentional (orienting) dysfunction in the ASD group, and serves as another distinguishing factor between ASD and other ESSENCE disorders. These findings are in line with most other current research which shows individuals with ASD, especially young children, are slower to fixate to social information than both
TD and developmentally delayed controls (for a meta analysis, see Guillon et al., 2014).

Additionally, correlations were found for all participants (regardless of diagnosis) between various social clinical measures (specifically VABS social, GAF, and FTF social competence) and a participant’s latency to social AOIs. This has important clinical implications, mainly in reinforcing the validity of these assessments, but also in outlining a potential early marker for ESSENCE (not restricted to ASD) since eye tracking can be used in infancy and social salience has been shown to be present in typically developing humans from early infancy and possibly even from before birth. It also indicates a common neurophysiological factor related to social orienting and alertness across diagnostic groups, which is a direction for future research.

Across diagnoses, the duration of fixations to the social AOIs correlated with the difference in pupil diameter between social and non-social images, suggesting that the longer a subject spent looking at social information, the greater their pupil dilation was in response. These results should be replicated and elaborated on before making assumptions, but could provide a possible direction for clinical intervention and insight to neurological underpinnings of social processing in ESSENCE disorders.

Interestingly, though there was a strong trend, no significant difference was found between groups for the total amount of time spent fixating to social information. This contradicts results from earlier studies. Perhaps this is due to methodological differences in the present study. For example, this study used static images, whereas some previous studies have used multimodal and/or dynamic social stimuli. This effect was seen by Anderson, Colombo, & Shaddy (2006), who used static social and non-social images and did not see decreased looking time to social images in patients with ASD. Baranek et al. (2013),
however, used a combination of social and non-social tactile, auditory, and visual stimuli and found decreased response to social stimuli in all categories. A meta-analysis of eye tracking studies by Guillon et al. (2014) revealed that in infants and young children, ASD patients showed decreased looking time to social information in general when the stimuli were dynamic and/or interactive, and showed no difference from controls when the stimuli were static images. With older children and adults, however, the results were not as distinguishable by stimulus type. An additional methodological consideration in the present study was the use of complex social scenes where faces occupied only 3-5% of the image; many existing studies have used only faces. Perhaps in a natural scene, social information is salient to individuals with ASD. Freeth et al. (2010) found that in static pictures of complex scenes, ASD adolescents spent a similar amount of time fixating to faces as TD adolescents. Fletcher-Watson et al. (2009) also found that ASD and TD adults spent a similar amount of time looking at faces within social scenes, and when person-present scenes were presented alongside person-absent scenes, both ASD and TD adults showed preference for the person-present scenes. There are several other studies with scenes, however, which show that ASD patients respond less to social information (Guillon et al., 2014), so this is perhaps an incomplete explanation. Another possible explanation for the difference in these results from previous findings could be that a majority of these stimuli depicted individuals with averted, rather than direct, gaze; i.e. the people in the images were not looking directly at the participant. Previous studies have used stimuli with direct gaze, and this could be the cause of ASD patients looking less at faces. Chawarska, Macari, & Shic (2012) showed that toddlers with ASD had similar responses to social scenes without eye contact and speech as TD and developmentally delayed controls, but when eye contact was added to the scenes the ASD group showed decreased
attention and responses to the social scenes. Hadjikhani et al. (2017) found that individuals with ASD show higher subcortical activation when a fearful stimulus face was looking directly at them instead of looking away, whereas TD controls showed more activation when the fearful stimulus looked away. This indicates that ASD individuals are more fearful of direct eye contact than of a potentially threatening object which is not visible. The difference in the present results from previous research could also be due to the absence of a typically developing control group; perhaps the ASD and non-ASD groups in this study, though similar to one another in fixation, would show differences from a typically developing control group. This would indicate a shared factor of social processing within ESSENCE.

In the present study, typically developing controls were not used, and instead the control group was composed of individuals with an ESSENCE diagnosis (excluding ASD). This was done because ASD is highly comorbid with other developmental disorders, and these disorders often display similar symptoms. In order to outline subtle, more exact differences and provide a clear picture of social processing in ASD, it is necessary to compare an ASD group with a group which has similar symptoms and even shares one or more diagnoses with the ASD individuals. For example, an individual may have diagnoses of ASD, ADHD, and epilepsy. In order to detect which symptoms and neurological factors are caused by ASD, it is necessary to have controls with diagnoses of epilepsy and/or ADHD; only in this way can one be sure to truly control for overlapping causes. The use of only a “typically developing” (which according to some suggestions in reality oftentimes are unusually super-normal [Kendler, 1990]) control group is problematic since it ignores the reality of ASD. For diagnostic and research purposes, it is valuable to utilize a neuropsychiatric control group to make ASD more detectable and outline specific markers and
The findings presented here are limited by the small spectrum of diagnoses in the non-ASD group. Though a few diagnoses under the ESSENCE umbrella were represented, ADHD was predominant. It would have been more advantageous to have a wider variety of disorders to allow for more sensitivity within the ASD categorization. The present study is also limited by a relatively small sample size. Though the ability to detect differences in ASD with a small sample size indicates an increased sensitivity to the categorization of ASD, a small sample size creates problems when attempting to generalize results to a larger population. In future studies and replications it is advantageous to address these limitations.

Together with other current research, these findings help to solidify and expand understanding of social processing in ASD and direct future questions. The current results are clearly more easily aligned with the reduced social motivation hypothesis than with the suggestion that the social difference and social disengagement that characterize ASD at the behavioral level reflect over-arousal during social processing. However, it is tempting to speculate whether perhaps the two leading hypotheses (gaze aversion and reduced social motivation) do not need to compete; the answer could lie in a unified approach. Taken together, the current available evidence supports the hypotheses that individuals with ASD are hyperaroused by at least eye contact (e.g., Dalton et al., 2005; Hadjikhani et al., 2017). At the same time in scene perception, as has been replicated here, social information appears to receive reduced priority, at least in the orienting phase (e.g., Klin et al., 2002; Senju & Johnson, 2009). Instead of teasing apart the two alternatives, future research should perhaps aim to combine them. Indeed the two ideas are not necessarily mutually exclusive; the answer could lie in the idea that individuals with ASD are inflicted with impaired social orient-
ing as a result of a learned response to hyperactivation when looking at social stimuli. In other words, people with ASD could experience hyperactivation in response to eye contact as infants, causing gaze aversion. As a result of gaze aversion from an early age, the social brain would not develop typically, leading to difficulties in orienting and attention in social settings. This could help to explain the discrepancies in results between various studies employing different research techniques and types of stimuli.

Studies on infants at high risk for ASD support this hypothesis. Jones and Klin (2013) found that 2-month-old infants who were later diagnosed with ASD spent the same amount of time looking at the eyes as 2-month-old TD infants. 6-month-old infants who were later diagnosed with ASD showed typical eye gaze patterns (Merin, Young, Ozonoff, & Rogers, 2007) and had the same looking times to the eyes and mouth as typically developing controls (Chawarska, Macari, & Shic, 2013). Elsabbagh et al. (2013) also found that infants with ASD and TD infants had similar attentional patterns to the eyes and mouth in a peek-a-boo sequence. Toddlers and young children with ASD generally show decreased attention to social stimuli, but studies using adolescents and adults with ASD have mixed results depending on the type and complexity of stimuli used (Guillon et al., 2014).

Moreover, eye tracking studies and other studies of social perception in ASD tend to be inconsistent depending on the type of stimuli used. Generally, studies using static images of faces with direct gaze find atypical gaze patterns, especially with regard to avoidance of the eyes (e.g., Guillon et al., 2014). Studies using more dynamic films seem to find that ASD participants look less often to faces than TD controls (e.g., Klin et al., 2002). When studies use static images of complex scenes, results are mixed with regard to the duration of social fixations but overall tend to find that ASD participants are slower to orient to
and fixate on the socially relevant information (e.g., Senju & Johnson, 2009). Consistent with these findings, the present study (which used static images of complex scenes featuring one person) found that ASD participants did not spend more time looking at social information than non-social information and took twice as long as the non-ASD participants to orient to the social AOIs. Pupil studies have shown larger tonic pupil diameter in ASD infants in comparison to TD controls (Anderson & Colombo, 2009) and decreased pupil diameter in ASD patients compared to TD and developmentally delayed controls in response to social information (Anderson & Colombo, 2006). This indicates a higher baseline autonomic arousal state, but a lack of attention to social information. These findings are in line with the results presented here, which show ASD participants had no difference in pupil diameter between social and non-social stimuli, whereas non-ASD participants had increased pupil diameter in response to social stimuli. fMRI studies have shown mixed results, with some showing lower relative subcortical arousal in response to social stimuli, and some showing increased arousal. This discrepancy can be explained by gaze, however; when subjects are specifically cued or instructed to look at the eyes, hyperactivation in the subcortical social brain is observed (Hadjikhani et al., 2017). When using SCR, results show increased skin conductance in response to social information (Joseph et al., 2008). Since orienting to and focusing on socially relevant information is essential for proper social cognition and behavior, dysfunction in the social brain experienced by individuals with ASD has cascading and lasting effects on their social functioning and cognition.

After reviewing the available current research, a pattern begins to emerge relating the observed behavioral and neurological responses of the ASD individuals: different types of stimuli evoke different responses. This makes it difficult to make a conclusion about social processing in general, as there is currently no
consensus on how participants respond to social information because the social components vary greatly across studies.

Figure 10: A summary of the aspects of stimuli which affect the behavioral and neurological response of ASD participants.

Figure 10 shows a summary and potential explanation of the varying results shown in current research. If the hypothesis presented here is correct, and ASD individuals experience hyperactivation and resulting attentional dysfunction, then different aspects of stimuli will evoke different engagement levels and brain arousal.

Visuo-spatial constraints refer to the technical aspects of a stimulus that can affect a participant’s ability to orient and arousal level. These constraints include the size of the social component in the image. For example, if the stimulus is a large scene where the face occupies a very small part of the image, ASD participants will be slower to orient attention to the face; if the stimulus is a large face only, the ASD participants have no choice but to see a face and thus will show hyperactivation and avoidance of the eye region (Guillon et al., 2014). Another visuo-spatial constraint is complexity. A stimulus could be a complex
social scene of which the social component is only one aspect of several other aspects; in this case, the stimulus is so to speak partly masked and may be less likely to evoke hyperactivation and instead will show slower orienting than if the stimulus shown has only social components and no additional components (Freeth et al., 2010; Klin et al., 2002). Complexity also includes how dynamic the stimulus is; whether it is an interactive video, a real person in the room with the participant, or a two-dimensional image can all elicit varying degrees of engagement and arousal (Baranek et al., 2013; von dem Hagen & Bright, 2017). The final visuo-spatial constraint is centrality, or where the social component of the stimulus is located. Again, in more complex scenes, if the social component is not centrally located the participants can be slower to orient to them, whereas a more centrally located social component may be more noticeable and evoke more arousal.

Intensity can be first determined by the emotionality of the stimulus. If the person in the stimulus is angry, happy, sad, or neutral, these can each have different interpretations by the participant (Lassalle et al. 2017). Gaze direction is another factor of a stimulus’ intensity. If the stimulus shows direct gaze, this can evoke hyperactivation because it is more engaging than indirect gaze (Chawarska, Macari, & Shic, 2012; Hadjikhani et al. 2017). Finally, the approach of a stimulus, or whether or not the person in the stimulus appears to be approaching the participant, can be interpreted as more or less threatening and thus can affect the participants’ engagement and arousal levels.

Attempting to draw conclusions about the same neurological processes from studies using different stimuli and methods can be problematic since there are so many aspects to stimuli which can alter a participant’s perception and response. This could be the reason there are so many conflicting results across many studies. In the future, it will be important to systematically use appropriate
stimuli for the hypotheses being tested, and to formally consider the effects of the various stimuli and the results of other studies when drawing conclusions.

6 Conclusion

This study aimed to examine scene scanning and pupil dilation in response to social information of an ESSENCE clinical group classified by either presence or absence of an ASD diagnosis to obtain a more sensitive and specific classification of ASD from other similar disorders. Across groups there was a correlation between all social psychiatric metrics and the latency to social information, which reinforces the validity of these evaluations and provides potential early clinical markers for ESSENCE. The ASD group took longer to fixate to the social information than the non-ASD group, indicating atypical orienting in ASD. The ASD group also showed no difference in fixation time between social and non-social information, though there was no significant difference in the duration of social fixation between the ASD and non-ASD groups. Additionally, the ASD group did not display increased pupil diameter in response to social information in comparison with the non-ASD group who showed larger pupil diameter when viewing social images than non-social images. Overall these results are in line with most current research and support aspects of the current hypotheses around social processing in ASD, leading to a new working hypothesis presented in this paper. The gaze aversion hypothesis suggests that individuals with ASD avoid looking at social stimuli because it causes subcortical hyperactivation and thus oversensitivity to social information. The social motivation hypothesis suggests that individuals with ASD do not look at social information because they have a failure to orient to and focus their attention on the social information, and do not experience intrinsic reward or motivation to look at it. Part of the reason there are so many conflicting results could be due to the vastly different
stimuli used across studies, the components of which can lead to different engagement and arousal levels in ASD participants. The current study suggests that rather than separating these two hypotheses research should aim to combine them. A new working hypothesis combining the results from this study and previous studies suggests that individuals with ASD experience hyperactivation and oversensitivity to eye contact, leading them to avoid eye contact early in development and causing atypical development of the social brain network. As a result, individuals with ASD show impairments in orienting and attention to social information. These developmental impairments and abnormalities may have a cascading effect on social processing which can last a lifetime.

7 Acknowledgements

I want to express my appreciation to my supervisors at the Gillberg Centre, Nouchine Hadjikhani, Jakob Åberg Johnels, and Martyna Galazka; their guidance was essential to the completion of my project. I would also like to thank Eva Billstedt, director of the Gillberg Centre, for allowing me the opportunity to complete my thesis in an ongoing project at the centre. My supervisor in Skövde, Judith Annett, was also helpful in reading drafts of my thesis paper.
8 References


Chawarska, K., Macari, S., & Shic, F. (2013). Decreased spontaneous attention to social scenes in 6-month-old infants later diagnosed with autism spectrum


Freeth, M., Chapman, P., Ropar, D., & Mitchell, P. (2010). Do gaze cues in complex scenes capture and direct the attention of high functioning adoles-


More eye contact, less amygdala activation. *Nature Scientific Reports, 8*, 3602.


fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry, 59*, 809-816.


Behavior Scales (2nd ed.) Circle Pines, MN: AGS.


