DISSERTATION | DOCTORAL THESIS

Titel | Title

The Neuropharmacology of Social Motivation in Autism: Integrating Opioid-Oxytocin Interactions, Naturalistic Eye Tracking and Computational Modeling

verfasst von | submitted by

Raimund Bühler BSc MSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 796 605 298
Dissertationsgebiet lt. Studienblatt | Field of study as it appears on the student record sheet:	Psychologie
Betreut von | Supervisor:	Univ.-Prof. Giorgia Silani Privatdoz. PhD

Table Of Contents

ACKNOWLEDGMENTS

Pursuing a PhD has been among the most challenging experiences of my life and it was a bumpy road at times. It made me recognize my limitations, but it also taught me the value of perseverance: to continue working even when the outcome can feel flawed and to stay committed in the face of adversity, all the while taking care of myself and those around me. Maintaining this balance is not always easy. What is much easier is to remember all the wonderful people who accompanied me along the way and supported me in so many ways. It would not have been possible without them. There are too many to name, I will do my best to give credit and I hope those not mentioned will forgive me.

First and foremost, I want to thank Giorgia. You have been the most supportive and reliable supervisor I could have wished for. Managing someone with high expectations can be a challenge, and you handled it with a calm and steady assurance. Your resilience, optimism and ability to move forward without panic are qualities I really admire. I have learned a great deal about leadership simply by being around you.

At the beginning, there were a few outstanding people who inspired me to even commit to a PhD and helped me to transition from my master studies. Most of all, Claudia, who was always willing to listen to my many questions and provided countless ways of guidance to the very end. Claudia, you have been an inspiration and a pillar of support, and I hope our paths will cross again someday. Similarly, Ana, thank you for always being there to answer my questions, and for all the balkan dance sessions. Seba and Emilio, both of you gave senior support whenever needed and were such wonderful lab colleagues.

Over time, some people left and others joined our lab. I remember meeting Asena when giving a talk on reinforcement learning and though the topic was quite challenging, you showed real engagement. You are an incredibly intelligent person and it is always a pleasure to talk to you. Thank you too for initiating the pasta lunches with me and for all the amazing Japan experiences. I will also fondly remember Yanick, who was one of the kindest and warmest people I ever met and so supportive in managing the studies. Yanick, I wish you all the best for your psychotherapy training. Also, thanks to all the others in the lab: Jakub, Alex, Kathi, Isa, Arya, Fiona, Ana and all the others. All of you are such wonderful colleagues and an invaluable part of my PhD journey. I also want to thank the people from SCAN unit for giving me shelter when I had no desk, you are just as much a part of my lab family.

The studies presented in this thesis would not have been possible without the tireless work of master students and technical support. Nina Sirec was the first to join my project and Libor Potocar went on to publish the second chapter of this thesis with me. A special thanks goes to all the students who took on the logistical challenge of running a four-session, within-subject pharmacological study: Ronja Lipp, who managed recruiting and social media channels, Nina Zeilinger, Chiara Vollmerhausen, Mehmet Uzunoglu, Rudolph Brozmann, Felix Puster and Felix Servus. Thank you all for your energy, patience, and dedication. The eye-tracking studies would not have been possible without the technical support of Ulrich Ansorge and Daniel Gugerell. The system for recruiting participants was set up with David Cserjan, who gave incredibly valuable input on how to implement a recruitment system. I'm also very grateful to Nace Mikus and Lei Zhang for brainstorming and discussing computational modeling approaches.

Lastly, I want to thank my parents and loving family, who provided a foundation of unconditional love and support. My mother, with her endless curiosity for literature, music, and crafts instilled in me a desire to understand and to create new things. My father remains a role model to me for his problem-solving skills, his unceasing pragmatic energy and his loving nature. These qualities have set the stage for my education, which will continue far beyond this thesis. And I want to thank Luise for going on this journey with me. I can barely express in my own words how grateful I am to you for being by my side. I will therefore conclude with the words of Daniel Johnston: True love will find you in the end.

GENERAL INTRODUCTION

During evolution, a diverse range of species has developed social behaviors that enable individuals to cooperate, foster complex relationships and build social hierarchies. Rather than enhancing the fitness of a single individual, social behaviors increase the fitness of the group as a whole, for example by facilitating access to resources that would otherwise be unattainable or by distributing them in an efficient way. Sociality and cooperation have therefore emerged as fundamental evolutionary strategies for survival. Accordingly, sociality is deeply engrained in the central nervous system of humans and other social species, warranting a long-standing scientific investigation into the neurobiological mechanisms that govern social behavior and how disruptions in these mechanisms can affect it.

One prominent theory proposes that the tendency to engage in affiliative social interactions relies on the same neural systems that drive other essential survival behaviors, most notably the *reward system*. This is known as the *common currency hypothesis* (Levy & Glimcher, 2012), which suggests that social rewards engage the same brain regions as other primary rewards, such as food and drink. Simply put, just as the reward system motivates us to seek nourishment, it also motivates us to engage in social behavior (Panksepp, 1979). One system most commonly implicated in the processing of primary rewards is the *opioid system,* which is thought to mediate the hedonic or pleasurable aspects of rewards (Berridge et al., 2009). A powerful illustration of the force this system exerts on behavior lies in opioid drugs, which can elicit intense feelings of pleasure but may also cause severe addiction, demonstrating how alterations in the opioid system can effectively derail adaptive behavior.

The common currency hypothesis is not without opponents: A competing view, the *social brain* perspective suggests that social stimuli are processed by dedicated neural systems and should be considered distinct (Dunbar, 1998; Frith, 2007). For example, in recent years a lot of research has been dedicated to the *oxytocin system*. Oxytocin, a peptide synthesized in the hypothalamus and released from the pituitary, was first investigated for its hormonal role in childbirth and lactation (Kendrick et al., 2017). However, research in monogamous prairie voles suggested broader effects in pair bonding (McGraw & Young, 2010) and later studies in humans indicated that oxytocin also affects trust (Kosfeld et al., 2005). The oxytocin system therefore appears to be a candidate for a largely dedicated neural substrate governing complex social behaviors (Kanat et al., 2014). However, the debate regarding the specificity of neural mechanisms involved in social behavior is ongoing and their separation in experimental research remains challenging.

Importantly, the opioid and oxytocin systems interact at the neural level, with opioids assumed to inhibit oxytocin release (Bicknell & Leng, 1982). The interaction between these systems suggests that modulating one may have cascading effects on the other, potentially influencing social behavior as well. While some research conducted in animals has demonstrated interactive effects on social attention (Dal Monte et al., 2017), this relationship remains largely unexplored in humans.

Furthermore, certain conditions are associated with social atypicalities, most notably autism spectrum disorder (ASD). ASD is a heterogeneous neurodevelopmental condition characterized by deficits in social communication and interaction, as well as restricted, repetitive behaviors and interests. Children with ASD often struggle to spontaneously share experiences, interpret nonverbal social cues, and develop, maintain, or understand social relationships (American Psychiatric Association, 2013). However, ASD is also associated with a set of cognitive strengths (Joseph et al., 2002): Individuals with ASD often demonstrate a detail-oriented cognitive style, an increased ability for pattern recognition and remarkable memory skills in areas aligned with their specialized interests. These qualities may contribute to success in domains where they are advantageous, such as mathematics,

science, music and visual arts. Moreover, ASD is now recognized as a dimensional construct, with autistic traits extending into the general population (Grzadzinski et al., 2013).

The underlying causes of social differences in ASD have been approached from various theoretical angles: the *social motivation hypothesis* proposes that individuals with ASD process social rewards differently, leading to a reduced motivation to engage in social interaction (Chevallier et al., 2012). According to this view, the fundamental difference lies in a diminished valuation of social stimuli, resulting in fewer social learning opportunities during development and, consequently, a lack of social skills later in life. However, a more general variation of cognitive processes with downstream effects on social behavior is also plausible. For example, according to the *predictive coding* account, ASD can be framed as a condition of *learning*, which in turn is strongly affected by reward processing (Van de Cruys et al., 2014). By acting as *reinforcers*, rewards can adaptively shape behavior, increasing the probability of obtaining them in the future. While this ties in closely with the motivational aspect, the predictive processing account usually assumes *domain-general* learning differences in ASD, which could potentially also explain symptoms in non-social domains, such as repetitive behavior or resistance to change. Nevertheless, these theoretical perspectives are not mutually exclusive and represent only a small subset of the broader etiological landscape in ASD.

In experimental research, social behavior has been measured in various ways, a prominent one being *eye tracking*. During social interactions, eye contact is among the most crucial social signals, conveying engagement, turn-taking and emotions. In visual experiments, social cues such as faces and eyes are highly salient, a phenomenon studied in the field of *social attention*. In these studies, attention to the eye region is often found to be reduced in ASD (Chita-Tegmark, 2016). However, due to technological restrictions, many studies have relied on static stimuli, such as pictures of faces displayed on screens, which lack the complexities of real social interactions (Chevallier et al., 2015). Recent advances in eye tracking and face detection via computer vision algorithms now enable more naturalistic experiments and many researchers advocate for studies which better approximate real social behavior (Schilbach et al., 2013).

While eye-tracking research provides valuable insights into social attention, it does not address whether underlying cognitive mechanisms are driven by uniquely social or domain-general processes. A promising approach to this problem is to frame social differences in ASD as emerging from atypical prediction and learning, as proposed by predictive coding theories. Probabilistic models of behavior, often referred to as *computational models* or *reinforcement learning models*, enable the estimation of latent parameters, offering precise insight into which components of learning are affected. Comparing learning across social and non-social contexts makes it possible to disentangle domain-specific from domain-general alterations in reward-based learning, thereby offering a critical evaluation of the social motivation hypothesis.

In summary, interactive pharmacological experiments offer a promising approach to elucidate the neurobiological systems that govern social behavior and how they may be altered in ASD. While extensive research has focused on the isolated effects of opioids and oxytocin, their interactive effects are less explored in humans. Furthermore, social attention has primarily been studied in screen-based experimental settings, with studies in naturalistic contexts still rare. Moreover, it remains unresolved whether differences in social attention reflect domain-specific mechanisms or broader alterations in learning and reward processing. The overarching goal of this thesis is therefore to investigate how the opioid and oxytocin systems interact to shape social attention

in ASD within a naturalistic setting, and to apply computational modeling to explore underlying mechanisms of social learning in autism.

In the following, I will provide a brief overview of the opioid and oxytocin systems and their interactions, with a particular focus on their roles in social attention and learning. I will then review current methodological approaches to studying social attention in ASD, comparing screen-based paradigms with more naturalistic designs and highlighting the potential of mobile eye tracking and computer vision for capturing real-life social gaze behavior. Lastly, I will briefly introduce the predictive coding framework, outline its connections to computational models of reinforcement learning and discuss how such models can disentangle domain-specific from domain-general learning differences in ASD. Together, these chapters establish the theoretical background and methodological approach guiding the empirical work presented in this thesis.

1 Pharmacological Modulation of Social Attention

The Opioid System and Social Reward Processing

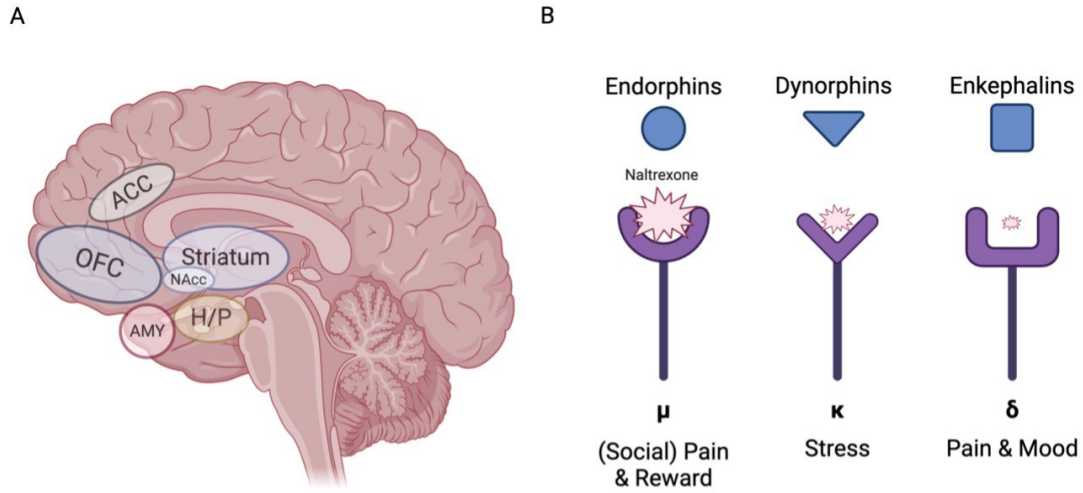

Figure 1 – The Opioid System. **A** Key brain regions: the orbitofrontal cortex (OFC), striatum and nucleus accumbens (NAcc) are highlighted for their roles in reward processing, while the ACC and insula are implicated in pain perception. The amygdala (AMY) is associated with the processing of emotion, stress, and salience and the hypothalamus/pituitary (H/P) is shown due to its role in stress regulation and oxytocin-opioid interactions. **B** Three major opioid receptor subtypes: mu (μ), kappa (κ), and delta (δ), along with their corresponding ligands: endorphins, dynorphins and enkephalins. Note that receptor functions are not exhaustive and may overlap. Naltrexone, an opioid antagonist, preferentially blocks μ-opioid receptors and to a lesser extent κ- and δ-receptors. Abbreviations: OFC, orbitofrontal cortex; NAcc, nucleus accumbens; ACC, anterior cingulate cortex; AMY, amygdala; H/P, hypothalamus/pituitary. Created in BioRender.com

The endogenous opioid system is comprised of three main peptides: endorphins, dynorphins and enkephalins. These peptides, composed of short chains of amino acids, act on three main opioid receptor subtypes: mu (μ), kappa (κ), and delta (δ), each of which is associated with distinct but overlapping functions (Waldhoer et al., 2004). Opioids were first investigated for their role in pain and euphoria (Brownstein, 1993). However, beyond its role in analgesia, the μ-opioid system is also strongly implicated in social behavior, influencing pair bonding (Burkett et al., 2011), social reward (Trezza et al., 2010), social pain (Eisenberger, 2012), pleasurable touch (Ellingsen et al., 2016) and attention to faces (Chelnokova et al., 2014, 2016). The κ-receptor, in contrast, is linked to stress modulation and dysphoria (Bruchas et al., 2010; Knoll et al., 2011), while the δ- receptor has been implicated in mood regulation, but also contributes to analgesia (Chu Sin Chung & Kieffer, 2013). However, it is important to note that experimental separation of receptor functions is challenging, as pharmacological agents are usually non-specific (Valentino & Volkow, 2018).

Opioid peptides and their receptors are expressed throughout the body, with the highest concentrations in the central and peripheral nervous systems, but also in the gastrointestinal tract (Machin & Dunbar, 2011). Within the central nervous system, specific opioidergic regions have been the focus of extensive research: the orbitofrontal cortex (OFC) and striatum, particularly the nucleus accumbens (NAcc), are central to reward processing, pleasure, and addiction (Berridge & Kringelbach, 2015). The anterior cingulate cortex (ACC) and insula are implicated in pain perception (Eisenberger, 2012), while the amygdala plays a key role in salience (Mahler & Berridge, 2009), and, along with the hypothalamus, pituitary and adrenal glands (HPA-axis), in stress responses (Knoll et al., 2011). Notably, the hypothalamus is also an important site of opioid-oxytocin interactions (Bicknell & Leng, 1982; Vandesande & Dierickx, 1975). Opioid agonists mimic the function of endogenous opioids by activating opioid receptors, whereas antagonists, such as naltrexone, block receptor activity and suppress endogenous opioid signaling with distinct profiles across receptor subtypes. For example, naltrexone exhibits the highest affinity for the μ- receptor, with lower affinity for κ- and lowest for δ-receptors (Trøstheim et al., 2022; see Figure 1B).

One of the earliest theories linking opioids to social behavior is the brain opioid theory of social attachment (BOTSA) (Panksepp, 1979). This theory was originally based on findings in rodents, where social isolation produced distress behaviors analogous to opioid withdrawal (Machin & Dunbar, 2011). Furthermore, morphine administration reduced separation distress vocalizations in rodents, chicks, and monkeys (Kalin et al., 1988), whereas opioid receptor blockade increased social seeking behavior (Fabre-Nys et al., 1982; Keverne et al., 1989). Based on these and further observations, Panksepp proposed that endogenous opioids mediate the rewarding aspects of social contact, and that social withdrawal or isolation results in an opioid-deficient state, thereby motivating social seeking to restore opioid tone (Panksepp, 1979). Consequently, the theory suggested that individuals with ASD might exhibit a permanently elevated opioid tone. This would in turn result in reduced motivation to seek social interactions, which could potentially be mitigated by opioid antagonists such as naltrexone. However, even in the early days, findings were not always consistent. For example, Martel et al., 1993 reported that naltrexone administration *reduced* maternal attachment and social grooming in rhesus monkeys, contrary to BOTSA predictions.

Nevertheless, BOTSA continued to be an influential theoretical framework and later studies supported the involvement of opioids in social attachment by demonstrating deficits reminiscent of ASD in μ-receptor knockout mice (Moles et al., 2004). In humans, positron emission tomography (PET) studies have shown that social touch and social laughter increase μ-opioid receptor activation in regions such as the striatum, insula, as well as frontal and cingulate cortices (Manninen et al., 2017; Nummenmaa et al., 2016). Genetic studies in humans also found that the OPRM1 A118G polymorphism, associated with increased μ-opioid receptor sensitivity, is related to stronger affiliative tendencies and heightened responses to social stimuli (Way et al., 2009). A growing number of more recent behavioral studies have further demonstrated the involvement of the endogenous opioid system in human social behavior across a diverse range of measures, including attention to the eye region (Chelnokova et al., 2016; G. Løseth et al., 2024).

Despite the evidence in support of BOTSA predictions, trials of naltrexone in children with autism spectrum disorder (ASD) have yielded mixed findings (for a review see Roy et al., 2015). While many studies report that naltrexone reduced hyperactivity and irritability, only a subset of studies reported improvements in social behaviors, such as increased eye contact and social initiation (Leboyer et al., 1992; P. G. Williams et al.,

2001). Therefore, current evidence remains insufficient to support the clinical use of naltrexone for addressing social symptoms in ASD.

Taken together, there is considerable evidence supporting the involvement of the μ-opioid system in social behavior (G. Løseth et al., 2024; Putnam & Chang, 2022b). However, the precise underlying mechanism remains unclear and evidence regarding the effects of opioid antagonists such as naltrexone is conflicting (Manduca et al., 2021). Specifically, it is not fully understood whether naltrexone increases or reduces sociality, with the effect likely dependent on context (Loseth et al., 2014; Pellissier et al., 2018). While in certain animal studies naltrexone increases sociality in line with BOTSA predictions (Keverne et al., 1989), in others it can also reduce it (Martel et al., 1993). In controlled laboratory settings with healthy human participants, naltrexone has been found to reduce social attention (Chelnokova et al., 2016), whereas some studies in children with ASD suggest increased eye contact (P. G. Williams et al., 2001). Some of these contradicting effects may be reconciled by accounting for the fact that the opioid system does not act in isolation: opioids interact closely with other neuromodulatory systems, particularly the oxytocin system (Bicknell et al., 1988; Bicknell & Leng, 1982; Shibuki et al., 1988).

The Oxytocin System and Its Role in Social Bonding

Oxytocin is a nonapeptide consisting of nine amino acids linked by peptide bonds, acting both centrally as a neuromodulator and peripherally as a hormone. When first synthesized, it was primarily known for eliciting uterine contractions and stimulating milk ejection (du Vigneaud et al., 1954). It was therefore employed in the clinical management of labor, a practice that remains standard today (Hermesch et al., 2024). Structurally, oxytocin is similar to vasopressin, another peptide associated with homeostatic functions, primarily water balance and thirst (Perisic et al., 2024). Both peptides are synthesized in the hypothalamus and transported to the pituitary, from which they are released into systemic circulation and exert their hormonal effects (Perisic et al., 2024). However, oxytocin also acts centrally as a neuromodulator, meaning rather than transmitting information itself, it modulates other neural systems (Bakos et al., 2018; Grinevich & Neumann, 2021).

The first observations suggesting effects of oxytocin beyond lactation and birth were gained in prairie voles, where a distinct pattern of oxytocin receptor density in certain brain regions was found for monogamous voles compared to non-monogamous voles, suggesting effects of oxytocin in pair bonding (Insel & Shapiro, 1992; Young & Wang, 2004). Further research in this species demonstrated that oxytocin also facilitates the formation of partner preferences by acting on receptors in the nucleus accumbens (Liu & Wang, 2003; J. R. Williams et al., 1994) and mitigates the behavioral consequences of social isolation, reducing depressive-like symptoms following separation from a bonded partner (Grippo et al., 2009).

In non-human primates, oxytocin enhances attention to socially salient cues, such as the eyes of conspecifics (Dal Monte et al., 2017). It also reduces social vigilance and attenuates attention to negative expressions, thereby promoting social approach (Ebitz et al., 2013; Parr et al., 2013). Lastly, it exerts effects on learning, amplifying vicarious reinforcement, meaning it increases learning from others' outcomes (Chang et al., 2012).

In humans, seminal evidence for oxytocin's role in social behavior came from a neuroeconomic study showing that intranasal oxytocin increases interpersonal trust in an economic game, marking a shift toward investigating oxytocin as a modulator of human social cognition (Kosfeld et al., 2005). Around the same time, oxytocin was shown to reduce amygdala responses to threatening stimuli (Kirsch et al., 2005), supporting a

potential anxiolytic effect. These early findings in humans catalyzed a surge of interest in oxytocin's social effects, including enhanced recognition of emotional faces (Di Simplicio et al., 2009), increased empathic accuracy (Bartz et al., 2010), and modulation of social memory (Rimmele et al., 2009).

More recent research highlights oxytocin's role in social attention: intranasal administration increases gaze to the eye region of faces (Auyeung et al., 2015; Guastella et al., 2008), enhances salience of social stimuli as captured by pupil dilation (Prehn et al., 2013), and modulates neural responses in the amygdala and insula to emotional cues (Striepens et al., 2012). In the domain of learning and reward-processing, oxytocin enhances reinforcement learning from social feedback (Hu et al., 2015; Hurlemann et al., 2010) and increases activity in reward-related areas during affective touch and face viewing (Scheele et al., 2013, 2014).

These findings have prompted growing interest in the therapeutic potential of oxytocin for neurodevelopmental conditions such as ASD. Early genetic studies identified reduced plasma oxytocin levels in autistic children (Modahl et al., 1998) and linked variation in the oxytocin receptor gene (OXTR) to social-communicative impairments and ASD risk (S. G. Gregory et al., 2009; LoParo & Waldman, 2015). Building on this foundation, initial behavioral studies demonstrated that intranasal oxytocin increased gaze to the eye region, improved emotion recognition, and enhanced social learning in individuals with ASD (Andari et al., 2010; Guastella et al., 2010). Although clinical trials have since reported mixed efficacy, some studies have shown improvements in social responsiveness and reductions in repetitive behaviors (Parker et al., 2017; Watanabe et al., 2015; Yatawara et al., 2016).

Despite promising initial findings, the field has since entered a period of critical reevaluation: Several larger-scale randomized controlled trials have failed to replicate earlier effects of intranasal oxytocin and meta-analyses generally have found limited evidence for a reliable effect on social functioning in ASD (Guastella et al., 2015; Ooi et al., 2016). As with naltrexone, the evidence does not yet support clinical use of oxytocin in ASD treatment (Guastella & Hickie, 2016). Nevertheless, the involvement of oxytocin in social processes can be regarded as well-established and, mirroring opioid research, mixed findings can potentially be attributed to contextual factors, prompting research into systems that may interact with oxytocin (Alvares et al., 2017; Putnam & Chang, 2022c; Quintana, 2018; Quintana et al., 2021).

Interactions Between Oxytocin, Naltrexone, and Social Gaze

A major impulse to consider oxytocin's interactions with other neural systems arose from the work of Steve Chang and colleagues, which emphasized interactions with opioids, but also with other neurotransmitters such as dopamine, serotonin and endocannabinoids (Putnam & Chang, 2022b, 2022c). Some of the first evidence for interactive effects on the cellular level came from a series of experiments by Bicknell, who demonstrated that naloxone administration significantly increased electrically stimulated release of oxytocin from the pituitary (Bicknell & Leng, 1982). Subsequent single-cell recordings revealed that naloxone increased the firing rates of oxytocin-releasing, but not vasopressin-releasing neurons (Bicknell et al., 1988), while morphine inhibited them (Brown et al., 2000). A corresponding pattern was found for morphine withdrawal, which selectively increased firing rates of oxytocin neurons (Brown et al., 2000), supporting the notion that endogenous opioids exert an inhibitory effect on oxytocin-releasing neurons.

Few studies have so far examined these interactive effects at the behavioral level (Putnam & Chang, 2022b). Seminal evidence came from an experiment conducted in macaque monkeys, which demonstrated increased fixations to the eye region of conspecifics after administration of both naloxone and oxytocin (Dal

Monte et al., 2017). These effects were strongest following socially salient events such as mutual eye contact or shared reward and were supported by genetic evidence of μ- and κ-opioid receptor gene colocalization with oxytocin receptor genes. More recently, a pharmacological study in mice demonstrated that systemic naloxone and central μ-opioid receptor blockade enhanced the anxiolytic-like effects of oxytocin, whereas κ-opioid receptor blockade attenuated them (Nisbett et al., 2024). In summary, these findings suggest that opioid antagonism may amplify oxytocin's social effects by removing endogenous opioid-mediated inhibition, potentially involving the κ-opioid system as well as the μ-opioid system.

Social Stress, Cortisol and Social Attention

While oxytocin and opioids exert well-established effects on social behavior, it is important consider the role of stress, particularly social stress, as a critical factor in the interplay between the opioid system and oxytocin (Beery & Kaufer, 2015; Sandi & Haller, 2015). Social stress responses are mediated by the hypothalamic-pituitary-adrenal (HPA) axis and its downstream release of cortisol (Kudielka et al., 2007). The κ-opioid system plays an important role in regulating stress responses (Knoll & Carlezon, 2010), with antagonism of κ-receptors leading to pro-social effects in animal models. For example, κ-opioid antagonism increases social play in rats (Vanderschuren et al., 1995) and reduces stress responses following social defeat (McLaughlin et al., 2006). In humans, increased social attention following naltrexone administration may be related to the inhibition of κ-opioid receptors (Wardle et al., 2016). Naltrexone and oxytocin exert opposing effects on cortisol: While naltrexone is known to activate the hypothalamic-pituitary-adrenal (HPA) axis, eliciting a cortisol response (King, 2002), oxytocin has been associated with stress-buffering effects, inhibiting cortisol release particularly in socially supportive contexts (Heinrichs et al., 2003). These opposing effects make it difficult to predict their combined effect on cortisol release.

However, in regards to social attention, cortisol responses to acute social stress have been directly linked to gaze behavior: In a virtual adaptation of the Trier Social Stress Test (TSST), individuals with a robust cortisol response (cortisol responders) exhibited less gaze fixation on socially relevant stimuli compared to non-responders (Vatheuer et al., 2021). In this context, reduced social gaze was significantly associated with an increased cortisol response, suggesting that stress-induced cortisol release may promote gaze avoidance. These findings are consistent with the broader literature indicating that stress and social anxiety contribute to gaze aversion as a coping strategy, a pattern that may also be relevant for ASD (N. J. Gregory et al., 2019). These findings implicate the HPA axis, stress responses and κ-opioid receptors as critical components in the modulation of social gaze.

2 Experimental Approaches to Social Attention in ASD

Due to its evolutionary importance, socially relevant information, such as facial expressions or the eye region of others, is highly salient, meaning it is prioritized during perception and sensory processing (Emery, 2000). This phenomenon, referred to as social attention, is associated with a distinct neural signature involving regions such as the amygdala (Nummenmaa & Calder, 2009). While the construct is generally considered multimodal, it is often operationalized in the visual domain using eye tracking (Valtakari et al., 2021). The most common measure is the allocation of attention to faces, particularly the eye region. Furthermore, even though social attention is considered a general mechanism of social cognition, it has mostly been studied in clinical contexts, especially in ASD and social anxiety disorder (SAD) (Chevallier et al., 2015; Günther et al., 2021). Across studies, individuals with ASD reliably show reduced fixations to the eye region. For example, in a meta-analysis Frazier et al., 2017 demonstrated that reduced gaze to eyes is one of the most consistent behavioral markers of ASD, showing moderate to large effect sizes across paradigms. Similarly, Chita-Tegmark, 2016 demonstrated that individuals with ASD show decreased attention to social stimuli across a range of task types. However, these findings are closely tied to the methodology of eye tracking, which has a long and evolving history. Understanding how social gaze is measured and how different experimental setups may influence observed effects is essential for gaining a complete picture of social attention differences in ASD.

Eye-Tracking Methodology

The investigation of eye movements dates back over at least a century with early experiments beginning in the field of psychophysics (Wade & Tatler, 2005). Today, eye-tracking experiments can be conducted in a variety of ways, depending on the research question and constraints of the study design (Nyström et al., 2025). Experimental setups typically fall into one of two categories: *screen-based*, which include head-restricted systems that stabilize the participant's head using a chin or forehead rest, and head-boxed systems that allow limited head movement within a defined tracking area; and *mobile*, head-free setups using head-worn eye trackers or eye tracking glasses (see Figure 2A). Notably, the majority of findings reporting reduced attention to social stimuli in ASD have been derived from screen-based paradigms, which, while well controlled and precise, lack ecological validity and fail to capture the complexities of real-world social interactions (Freeth et al., 2013; Risko et al., 2016) Furthermore, even within screen-based settings, social gaze in individuals with ASD varies depending on the nature of stimuli: dynamic, interactive scenes can differentiate individuals with ASD from typically developing participants, whereas static or non-interactive stimuli often fail to do so (Chevallier et al., 2015).

Over the past two decades, miniaturization of computers and cameras have enabled head-worn eye tracking glasses, which allow studying eye gaze in naturalistic environments. An early example of a study employing head-worn eye trackers was conducted by Land et al., 1999, who investigated eye movements during natural daily activities such as tea- and sandwich-making. In the context of ASD research, one study of particular relevance used eye-tracking glasses to compare gaze behavior across three social contexts: watching a video, interacting in a video-call and interacting in a face-to-face conversation (Cañigueral et al., 2021).

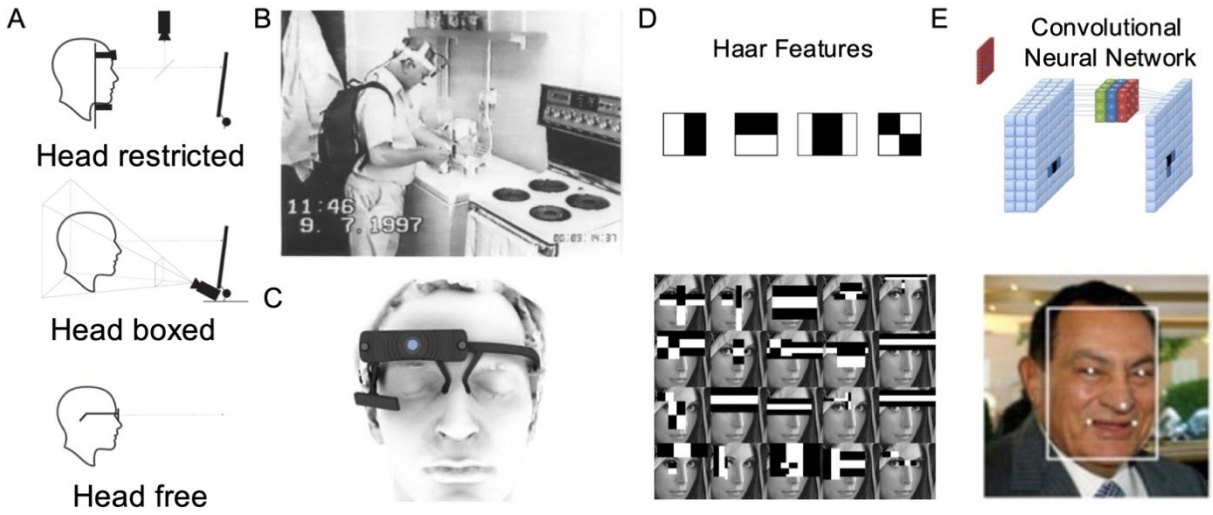

Figure 2 – Overview of Eye Tracking Setups and Face Detection Methods. A Three common eye tracking setups: two screen-based (head-restricted and head-boxed) and one mobile, head-free setup using head-worn eye tracking glasses. Adapted from Nyström et al., 2025 **B** Frame from an early head-mounted setup used in Land et al., 1999 in a naturalistic eye-tracking study. **C** Pupil Labs Core headset (2014 version), equipped with an infrared eye camera and a front-facing world camera. **D** Illustration of the Viola-Jones face detection method, using Haar-like features: small rectangular masks capturing edge, line, and center-surround differences. The lower panel demonstrates the application of these features to a face. **E** Schematic of a convolutional neural network (CNN) architecture, illustrating convolutional layers and kernel. The bottom panel shows a bounding box and extracted facial landmarks (eyes, nose, corners of mouth) returned by a MTCNN (Multi-task Cascaded Convolutional Neural Network). Licensing and attribution details are provided in the footnote.

Results showed that gaze patterns of autistic and control groups did not differ: both groups modulated eye gaze and facial expressions similarly across conditions, looking less at their interaction partner and producing more facial expressions in the face-to-face conversation. Another study used eye tracking glasses during structured, face-to-face interviews with neurotypical participants varying in autistic traits (Vabalas & Freeth, 2016). While there was no relationship between autistic traits and the total duration of gaze to the partner or their face, individuals with higher autistic traits produced fewer and shorter saccades, indicating reduced visual exploration. These findings underscore that group differences in social gaze may depend to a large extent on experimental settings, with more naturalistic studies needed to elucidate social gaze dynamics in real-life situations (Schilbach, 2015).

The lack of such studies can potentially be attributed to additional challenges introduced by face-to-face interactions with head-worn eye trackers. This is primarily because areas of interest (AOIs), such as the face or mouth, are not static, but move in the participants visual field due to head and body movement of both the participant and the interaction partner. Researchers can deal with this problem in a number of ways: One approach involves simplifying the interaction, for instance via a video-call setup where the conversational partner remains fixed within a screen-based frame (e.g. Auyeung et al., 2015). Alternatively, AOIs can be manually coded frame-by-frame (e.g. Vabalas & Freeth, 2016), though this approach is time-consuming and effortful. More recently, advances in computer vision have enabled the automatic detection of faces and facial landmarks using face detection algorithms.

Face Detection Algorithms

An early and influential face detection algorithm is the Haar Cascade, introduced by Viola & Jones, 2001. The method relies on a set of Haar-like features (named after Hungarian mathematician Alfréd Haar) which consist of simple rectangular patterns designed to capture intensity contrasts in the image (see Figure 2). These features function much like edge detectors, quantifying the difference in summed pixel values between adjacent regions, such as the contrast between the eyes and cheeks, or the bridge of the nose and surrounding areas. This approach was used in some naturalistic studies of social attention in ASD (Noris et al., 2012). The authors used a multi-view Haar cascade system to analyze video from head-mounted cameras and found that children with autism spectrum disorder (ASD) spent significantly less time fixating on faces in naturalistic, free-play interactions compared to typically developing controls.

Algorithms like Haar cascades laid the groundwork for face-detection, but they were often less applicable for real-life scenarios, as they struggled with variations in lighting, pose and occlusion. With the advent of neural networks in the last decade, face detection has been dramatically improved using multi-task cascaded convolutional neural networks (MTCNNs) (Zhang et al., 2016). These are deep-learning approaches based on artificial neural networks with many layers sequentially arranged in a "deep" structure. Each layer in the network transforms the input data into increasingly abstract representations. In convolutional neural networks (CNNs) each layer performs a mathematical operation called a convolution: it applies a small, trainable filter (or kernel) of a specific size across the input image to detect specific local patterns, such as edges. Early layers might only detect simple features like contours, while deeper layers combine these to recognize more complex structures, like facial features or entire faces. Through this hierarchical processing, CNNs learn to localize faces and facial landmarks, which can then be used as areas of interest in studies of social attention.

In MTCNNs, neural networks are organized in a cascade, meaning that each stage progressively refines the output of the previous one. First, a potential face and its facial regions are proposed, then refined by a separate network and lastly, the final bounding boxes and facial landmarks are returned by the output network (Zhang et al., 2016). When trained on diverse datasets, this greatly enhances the accuracy and performance even in challenging, real-world conditions. The potential of using neural networks for automatic AOI detection in the context of social attention research is exemplified by a study, which directly compared AOIs generated via automatic face detection and manual coding. This study used the OpenPose model, a convolutional neural network for full-body pose estimation (Jongerius et al., 2021). While not a dedicated face detection model, OpenPose can detect facial landmarks such as eyes and nose (Cao et al., 2021). Results indicated very satisfactory interrater reliability of 0.85 to 0.98 and a normalized difference between methods of less than 2%, highlighting the emerging potential of face detection algorithms to advance studies of social attention.

3 Computational Models of Social Learning in ASD

Eye tracking provides valuable insights into *where* individuals with ASD direct their gaze, but not *why* they look there. As such, it captures the behavioral surface of social interaction without revealing the underlying cognitive mechanisms which give rise to social differences. For example, individuals with ASD may fixate on different locations during interactions because of how they interpret, predict, or value social information. To explore these underlying mechanisms, we can make use of computational models of learning, which formalize how individuals update beliefs and assign value in response to social and non-social feedback.

Computational models are ultimately rooted in the behaviorist tradition, which conceptualizes learning as a process by which individuals adapt their behavior in response to experience. In cognitive neuroscience and psychology, learning is often understood as a process of updating internal models, mental representations of how the world works, based on feedback from the environment. Computational models of learning make this process explicit: they mathematically describe how incoming information interacts with prior expectations to shape future behavior. By quantifying learning dynamics, such as how quickly beliefs are updated or how strongly outcomes are weighted, these models offer a framework for understanding individual differences in perception, cognition and action.

Although ASD is considered primarily a social condition, theoretical accounts that frame it as a learning condition argue that while it manifests in social contexts, the underlying cognitive differences are not necessarily confined to the social domain. Domain-general accounts, such as the influential *predictive coding* perspective, suggest widespread differences in how individuals with ASD update beliefs and process uncertainty. In contrast, domain-specific perspectives, sometimes linked to the social motivation hypothesis, focus on learning impairments that are restricted to social contexts (Ewbank et al., 2017; Kinard et al., 2020; von der Lühe et al., 2016). In both frameworks, applying computational models of learning offers a powerful tool for uncovering the specific cognitive processes which may be altered in ASD.

Predictive Processing and Bayesian Learning

The mathematical foundation for current computational models of learning was developed within the predictive coding account (Lawson et al., 2014; Mathys et al., 2011; Pellicano & Burr, 2012). This framework builds on the principles of Bayesian learning and proposes that the brain functions essentially as a prediction machine: it continuously generates expectations about incoming sensory inputs and updates internal models to minimize the mismatch between predictions and outcomes, so-called *prediction errors*. At the heart of this framework lies Bayes theorem, a foundational law in probability theory that, if interpreted as a mechanism for learning in probabilistic environments, formalizes how beliefs are updated based on new evidence while taking into account prior expectations.

$$\overbrace{p(\,\text{belief} \mid \text{evidence}\,)}^{\text{posterior}} = \frac{\overbrace{p(\,\text{evidence} \mid \text{belief}\,)}^{\text{likelihood}} \cdot \overbrace{p(\text{belief})}^{\text{prior}}}{p(\text{evidence})} \tag{1}$$

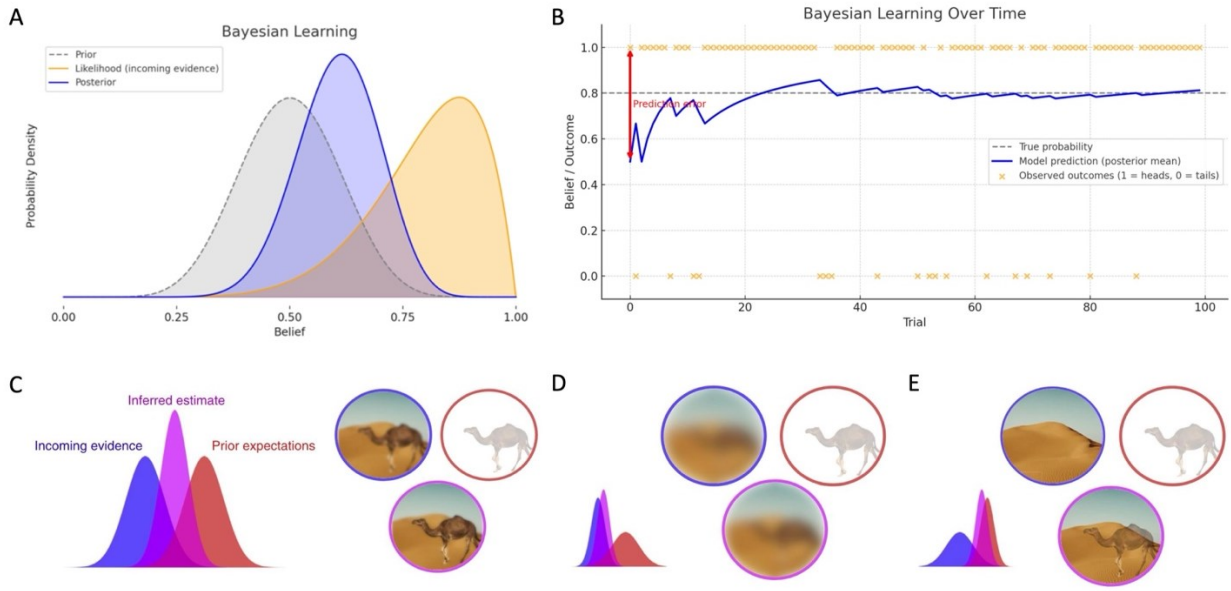

Figure 3 – Bayesian Learning in predictive coding and its alteration in neuropsychiatric conditions. A Bayesian belief updating involves combining a prior belief (gray) with new evidence (orange likelihood), resulting in a posterior distribution (blue) that shifts the belief towards the observed data. **B** Trial-by-trial simulation of Bayesian learning in a probabilistic coin-flipping task (true probability = 0.8). The model prediction (posterior mean; blue) gradually converges toward the true underlying probability (gray dashed line) as observations (orange dots) accumulate. A red arrow illustrates the prediction error, the difference between expected and observed outcome, on the first trial. **C** In typical inference, perceptual beliefs are formed by combining prior expectations (red) and sensory input (blue) in a precision-weighted manner, resulting in an inferred estimate (purple). **D** In ASD, the precision of sensory input is sometimes assumed to be abnormally high, leading to an over-reliance on immediate evidence and unstable perceptual beliefs. **E** Psychotic conditions such as schizophrenia may involve exaggerated precision of prior beliefs, which can dominate inference and result in perceptual experiences despite contradictory or absent sensory evidence. Panels C–E adapted from Yon & Frith, 2021. Precision and the Bayesian brain. Current Biology, 31(17), R1017–R1034, © 2021 Elsevier, with permission.

This principle is illustrated in Figure 3A, which visualizes how Bayesian learning combines prior beliefs with incoming evidence to form an updated belief, or posterior. In this process, the prior represents the learner's initial expectation, what is believed to be true before new data is observed. The likelihood reflects the sensory evidence, which may either confirm or challenge the prior belief. The resulting posterior represents the updated belief, integrating both sources of information. When prior and likelihood are in disagreement, the posterior shifts in the direction of the likelihood, weighted by their relative certainty or *precision*. The more precise (i.e., narrow) the prior or the likelihood, the more strongly it pulls the posterior toward itself. The denominator in Bayes' theorem, p(evidence), serves as a normalizing constant that ensures the posterior forms a proper probability distribution. While it is mathematically necessary for normalization, it plays no role in determining the shape of the posterior and is therefore often omitted in learning algorithms focused on proportional updates.

To make a concrete example, imagine learning about a biased coin with an 80% chance of landing on heads. Suppose a learner starts with no strong expectation about the outcome of the coin toss, an uninformative prior centered at 0.5, reflecting the belief that heads and tails are equally likely. Although the most appropriate distribution for modeling uncertainty about binary probabilities is the Beta distribution, here we approximate the learner's prior belief with a Gaussian distribution centered at 0.5 for illustration purposes.

$$p \sim \mathcal{N}(\mu = 0.5, \sigma^2) \quad (2)$$

The learner's prediction for a trial $\hat{\mu}_t$ corresponds to the expected value $\mathbb{E}$, which in the case of a Gaussian is equivalent to the mean μ of this distribution:

$$\hat{\mu}_t = \mathbb{E}[\mathrm{p_t}] \tag{3}$$

This prediction is then updated in light of new information. Suppose in the first trial, the coin lands on heads. At this point, we experience a critical element of the predictive coding framework: the *prediction error* δ. This captures the misalignment of our initial estimations and the outcome we observe. Mathematically, it is the difference between the outcome r_t and our prior expectation $\hat{\mu}_t$.

$$\delta_t = r_t - \hat{\mu}_t \tag{4}$$

We then update our expectation based on the prediction error. However, in bayesian frameworks this update is weighted by a *precision weight* α_t, which reflects how much confidence we place in the new observation relative to our prior belief.

$$\mu_{t+1} = \hat{\mu}_t + \alpha_t \cdot \delta_t \tag{5}$$

Precision is defined as the inverse of the variance, and in this context it applies to both components: the sensory input (i.e., the outcome of the coin toss) and the internal model (i.e., the prior belief about the probability of heads). When our prior is uncertain (i.e., has high variance), the prediction error exerts a stronger influence on the update. Conversely, if our prior belief is highly precise (has low variance), it dominates the update. The effective weight is thus determined by the relative precision of observation and belief. The following equations formalize the weighting of variances $\hat{\sigma}$ *and* precisions $\hat{\pi}$ in this context.

$$\alpha_t = \frac{1}{\frac{1}{\hat{\sigma}_{\mathrm{prior}}} + \hat{\sigma}_{\mathrm{sensory}}} \tag{6}$$

$$\pi_t = \frac{1}{\alpha_t} = \hat{\pi}_{\mathrm{prior}} + \frac{1}{\hat{\pi}_{\mathrm{sensory}}} \tag{7}$$

As the learner becomes familiar with the coin more often landing on heads, the prediction gradually shifts towards a higher probability of seeing heads (see Figure 3B). With each additional observation, the posterior distribution becomes more concentrated around the true underlying probability of 0.8, as the precision of the belief increases with accumulated evidence. However, the rate of belief updating and the stability of these beliefs can vary across individuals. A central feature of the predictive coding framework, particularly when applied to neurodevelopmental or psychiatric disorders, is the interplay between prior and sensory precision. Some individuals may rely heavily on sensory input, while others may overweight prior beliefs. The former is sometimes associated with ASD, the latter with schizophrenia and related psychotic disorders (Yon & Frith, 2021).

In ASD, the hypothesis derived from predictive coding suggests that sensory precision may be abnormally high, resulting in suboptimal learning and difficulties in extracting stable inferences from the environment. When incoming information is weighted too strongly, even minor deviations can cause disproportionate belief updates, resulting in sensory overload and an inability to stabilize predictions over time. This could potentially explain the detail-oriented cognitive style, resistance to change and social symptoms in ASD, as social environments are inherently complex and require flexible adaptation of expectations. In essence, posterior estimates remain overly close to sensory input, impeding the ability to make adaptive inferences about the environment (see Figure 3D). In contrast, psychotic disorders such as schizophrenia may be characterized by an overweighting of prior beliefs to the extent that they dominate perception. In extreme cases, this can result in

hallucinations: perceptual experiences which persist despite the absence of sensory evidence (see Figure 3E). The important insight for ASD is that the critical difference according to the predictive coding account may lie in the precision weighting of prediction errors, where sensory experience may be overweighted compared to prior beliefs. This imbalance has been described as the "*weak priors*" hypothesis.

Rescorla-Wagner models

As pointed out by Mathys et al., 2011 and Vossel et al., 2014, the belief updating equation used in predictive coding models shares structural similarities with the classic Rescorla–Wagner reinforcement learning model (Rescorla & Wagner, 1972).

$$\underbrace{\mu_{t+1}}_{\text{prediction at } t+1} = \underbrace{\hat{\mu}_t}_{\text{prior prediction}} + \underbrace{\alpha}_{\text{learning rate}} \cdot \underbrace{\delta_t}_{\text{prediction error}} \tag{8}$$

The most important parameter of the Rescorla-Wagner model is the *learning rate*, which incorporates both sensory and prior precision and is equivalent to the precision weight in predictive coding. The main difference between Rescorla-Wagner reinforcement learning and Bayesian predictive coding frameworks lies in the nature of the learning rate: in Rescorla-Wagner models, the learning rate is typically treated as a fixed parameter specific to each individual, whereas in predictive coding, it is dynamic, adapting to trial-wise estimates of uncertainty. More recent accounts of predictive coding in ASD move beyond simple differences in learning rates and propose hierarchical frameworks of learning which take into account changes in the environment or *volatility* (i.e. the Hierarchical Gaussian Filter (HGF), Mathys et al., 2011). These accounts often suggest that differences in ASD may lie not necessarily in learning rates, but rather in estimating the volatility of the environment (Lawson et al., 2017). Nevertheless, even within fixed-rate models, differences in the weighting of sensory versus prior information would reflect in the learning rate: In the context of ASD, weak priors (reduced confidence in prior beliefs) would result in *higher learning rates*, leading to faster adaptation to incoming sensory input.

A variation of Rescorla-Wagner models, which is particularly relevant in the context of social motivation and reward processing introduces another parameter at the level of the prediction error: *reward sensitivity* ρ (Huys et al., 2013). This parameter has been employed in research on mood disorders to capture the subjective impact of rewards during learning. Conceptually, it scales the influence of the received reward on belief updating, allowing for individual differences in how reward signals drive learning.

$$\delta_t = \boldsymbol{\rho} r_t - Q_t \tag{9}$$

This distinction seems particularly relevant in the debate on domain-specific vs domain-general learning differences, as the motivational value assigned to social vs. non-social stimuli can be inferred from the reward sensitivity parameter. Although reward sensitivity and learning rate both modulate the update of predictions, they are theoretically separable (Huys et al., 2013). Specifically, reward sensitivity influences the asymptote, the level to which the expected reward probability converges, while the learning rate determines how quickly this asymptote is reached. In the context of social vs. non-social reward processing, a blunted response to social rewards would manifest in reduced reward sensitivity, but specifically for social contexts. Importantly, this alteration can occur independently of learning rate differences.

It is important to note that some studies have also challenged the notion that predictive processing in ASD can be characterized by increased learning rates. For example, Lieder et al., 2019 proposed that individuals with ASD exhibit a "*slow-updating*" profile, reflecting *reduced* weighting of recent experience. In support of this view, Vishne et al., 2021 found that individuals with ASD showed impaired sensorimotor synchronization in a

paced finger-tapping task. A second experiment with tempo-switching further revealed that individuals with ASD were slower to adapt to changes in external rhythm, confirming a reduced rate of updating in dynamic environments. The authors interpret these findings as consistent with a slow-updating hypothesis, in which recent sensory information is underweighted. These results suggest that, while the predictive processing framework offers a compelling theoretical account, the empirical evidence remains mixed, with the operationalization of predictive coding varying considerably across studies (for a review see Cannon et al., 2021).

Importantly, learning differences in ASD may arise at early stages of sensory processing (Cannon et al., 2021). Consequently, stimulus complexity and low-level visual features are important variables to control for in experimental designs aimed at isolating social learning mechanisms. *Point-light displays* (PLDs) offer a promising approach in this context. By depicting motion through dot markers placed at key anatomical points, PLDs retain dynamic social information while eliminating confounding visual features like texture or color. PLDs have been extensively used to study biological motion perception, often through full-body displays (for a review see Pavlova, 2012). In this context, processing biological motion has sometimes been found to be impaired in ASD, particularly when emotional or socially relevant content is involved. Deficits manifest in both the perceptual discrimination of human motion from scrambled stimuli (Blake et al., 2003) and in the spontaneous interpretation of emotions conveyed by body movements (Hubert et al., 2007). Beyond bodily motion, PLDs have also been used to examine facial expressions. For example, Keating et al., 2022 employed dynamic facial PLDs to investigate emotion recognition, and found that autistic participants were less accurate in identifying anger. These results highlight the potential of facial PLDs to examine social learning while minimizing confounding visual features.

4 Thesis Outline

Building on existing research on the neurobiology of social behavior and its alterations in ASD, this thesis aimed to investigate the mechanisms underlying social attention and social learning across three empirical studies. I applied a multi-method approach combining eye tracking, computational modeling, and psychopharmacology to examine the biological underpinnings and domain-specificity of social differences in ASD. The overarching goal was to understand how social differences in ASD arise and how they may be affected by interacting neurochemical systems. Specifically, this thesis addresses three gaps in the existing literature. First, while social attention has traditionally been investigated using highly controlled, screen-based paradigms, it remains unclear whether these findings generalize to real-world interactions. Second, the mechanisms underlying reduced social attention in ASD remain debated, particularly with respect to whether they reflect a domain-specific deficit in social motivation or a broader alteration in learning dynamics. Third, although the opioid and oxytocin systems have independently been implicated in modulating social behavior, their interactive effects have scarcely been explored in human studies.

In **Chapter 1**, I explored how naturalistic social attention relates to autistic traits in the general population. Using eye tracking glasses and automated face detection via multi-task convolutional neural networks (MTCNNs), participants' gaze behavior was recorded during a semi-structured face-to-face conversation. I hypothesized that higher autistic traits would be associated with reduced fixations to the eye region. This chapter establishes the methodological feasibility of combining naturalistic paradigms with automated gaze analysis and serves as a proof of concept for the experimental approach used in subsequent studies.

In **Chapter 2**, I examined the specificity of social learning differences in ASD by applying computational reinforcement learning models to social and non-social learning contexts. Participants learned reward contingencies from either social (facial point-light displays) or non-social (abstract point-light displays) stimuli, and their behavior was modeled using a Rescorla-Wagner learning framework. Parameters for learning rate and reward sensitivity were extracted to assess whether autistic traits are associated with general learning differences or a specific reduction in the value assigned to social stimuli. This chapter addresses ongoing debates about domain-specific versus domain-general learning mechanisms in ASD.

In **Chapter 3**, I investigated the effects of pharmacological modulation on social gaze behavior in a naturalistic setting. In a double-blind, placebo-controlled, within-subject design, participants with and without ASD received oxytocin, naltrexone, both drugs combined, or placebo across four sessions. Eye-tracking data was collected during a face-to-face conversation, and saliva samples were taken to assess cortisol responses. In line with animal research, I hypothesized that both naltrexone and oxytocin would increase social gaze, with their combined administration producing a stronger effect than either drug alone. This chapter integrates behavioral, pharmacological, and physiological measures to explore the pharmacological modulation of social attention and its effect in ASD.

Together, these three chapters investigate the mechanisms of social attention and learning in ASD across multiple levels, combining naturalistic behavior, computational modeling, and pharmacology. The findings address ongoing debates surrounding social motivation, predictive coding, and neurobiological theories in ASD, and highlight the potential of real-world paradigms for advancing social neuroscience.

Declaration of AI Use

During the writing of this work, I used generative AI (ChatGPT, model 4o), particularly for language editing purposes. The model was employed to improve grammar, clarity, and stylistic flow. All conceptual content, scientific reasoning, and interpretations go beyond the mere generative capabilities of large language models and reflect the author's original work and intellectual contributions. ChatGPT was not used to generate research ideas, design experiments, or formulate theoretical claims. All references and citations were verified by the author, who takes full responsibility for the accuracy of the content and the integrity of the scientific arguments presented.

Chapter 1 – Social Attention in the Wild - Fixations to the Eyes and Autistic Traits during a Naturalistic Interaction in a Healthy Sample

Raimund Buehler*[,1], Ulrich Ansorge[2,3,4], Giorgia Silani*[,1]

[1]Department of Clinical and Health Psychology, Faculty of Psychology, University of Vienna, Vienna, Austria

[2]Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, Vienna, Austria

[3]Vienna Cognitive Science Hub, University of Vienna, Vienna, Austria

[4]Research Platform Mediatized Lifeworlds, University of Vienna, Vienna, Austria

Corresponding authors:

Raimund Buehler
raimund.buehler@univie.ac.at
+43-1-4277-47241
Liebiggasse 5, 1010 Vienna, Austria

Giorgia Silani
giorgia.silani@univie.ac.at
+43-1-4277-47223
Liebiggasse 5, 1010 Vienna, Austria

Abstract

Attention to social stimuli is a key component of social behavior and facilitates the development of fundamental social skills. Studies investigating social attention in neurotypical or neurodiverse populations have often relied on screen-based experiments using static images or videos, which lack the sensory richness and reciprocity present in real-life social interactions. This can possibly be attributed to the challenges one encounters when creating naturalistic experiments, such as dealing with dynamically moving areas of interest (AOI's), which require either time-intensive manual coding or restraining of participants. Here, we present findings from a paradigm using unrestrained mobile eye-tracking and a face detection algorithm (MTCNN) to measure fixation rates during a semi-structured, face-to-face interview. Data from N = 62 healthy adult participants was analyzed for gaze behavior and related to participants' autistic traits. We observed a significant negative correlation between fixation rates on the eye region averaged over the entire interaction and scores on the autism spectrum quotient (AQ) ($r = -0.14$), indicating participants with high autistic traits were fixating less on the eye region. We also compared different types of interview questions (open vs. closed) to explore whether the reduction in fixation rates was more pronounced for specific time intervals during the interview. Lastly, we discuss both possibilities for extensions as well as limitations of the presented paradigm that could serve as inspiration for future research.

1 Introduction

At almost every moment of our life, we are surrounded by an overwhelming amount of information. Selectively attending to only the most relevant stimuli is vital for adaptive functioning and well-being. Social stimuli, such as faces, voices or bodily movements, carried a particular importance for human behavior during evolution[1] and are often preferentially attended to by typically developing individuals[2]. This bias towards social stimuli, present from early infancy onwards, facilitates social interaction and the acquisition of social skills across development[3,4].

In autism spectrum disorder (ASD), an overall reduction in attention to social stimuli has been observed[2], with the largest effect sizes found for the eye region[5]. This effect is present across children[6], adolescents[7] and adults[8]. Excess gaze to mouth has also sometimes been reported in children[9] and adults[7], however this failed to replicate in a number of later studies[10]. Presumed aberrations in visual social attention seem to generalize to the broader autism phenotype, operationalized by autistic trait questionnaires such as the autism spectrum quotient (AQ)[11–15], the social skills subscale of the AQ[16], the Social Responsiveness Scale (SRS)[17] or when examining parents of ASD individuals[18].

Notably, to investigate mechanisms of social attention, researchers have predominantly relied on screen-based experiments[19]. In these settings, static or dynamic stimuli are displayed, such as pictures of faces or videos of social scenes[20–22]. This has resulted in a wealth of knowledge about specific mechanisms of social attention such as gaze cueing[23], emotion recognition[24,25] and inferring mental states from the eye region of others[26]. Additionally, a bias for the eyes has been observed across a range of stimuli, i.e. pictures of faces[27], complex social scenes[28], and dynamic videos of social interactions[29].

However, several researchers have questioned the generalizability of findings generated in such settings to real-life social situations[30,31], pointing out important differences in results gained from screen-based and naturalistic experiments[16,21,29]. Importantly, self-reported difficulties and clinical assessments of eye contact occur during real, face-to-face social interactions[32,33]. This has led to calls for more experiments to investigate social attention "in the wild"[31,34,35], involving the actual presence of another person[12,16], with the possibility to study the temporal and reciprocal nature of social interactions.

One major challenge in creating naturalistic experiments are dynamically moving areas of interest (AOI's), which require time-consuming manual coding procedures[11,36] or restraining of participants. Some researchers have therefore opted to reduce the complexity of the interaction, i.e. by creating a video-call setup, to ensure static AOI's across the duration of the experiment[11,37]. Others have implemented computer vision algorithms to automate the detection of faces and facial landmarks[36,38–41]. Early studies have used comparatively simple face detection algorithms[41], however more sophisticated approaches involving convolutional neural networks (CNN's) trained for face detection are emerging[38], with promising results when compared to manual coding procedures[36,39].

Still, the implementation of computer vision algorithms remains challenging and most studies in naturalistic settings continue to use manual coding of AOI's[36]. The goal of this study was therefore to investigate the feasibility of using face detection algorithms in naturalistic experiments and to explore the effects of autistic traits on gaze behavior using unrestrained mobile eye tracking and computer vision assisted coding of AOI's. We also aimed to explore characteristics of gaze behavior over the timecourse of the interaction, specifically in response to different types of questions asked during a semi-structured interview[37,42].

2 Materials and methods

2.1 Participants

All participants gave written informed consent before participation. Experimental procedures were carried out in accordance with the Declaration of Helsinki and experimental protocols were approved by the ethics committee of the Medical University of Vienna (EK-1393/2017). Informed consent for the publication of identifiable images in an online open-access format was obtained from depicted members of the research team (Figure 1).

The total recruited sample consisted of N = 74 participants. 13 participants were psychology students, who received course credit for participation. 61 participants were not studying psychology and were recruited via another channel. These participants received 10€ for compensation. The sample size was based on an a priori power analyses using G*Power, assuming a medium effect size of r = 0.3 and a power of β = 0.8. This resulted in an estimated sample size of 64 participants, however we decided to include slightly more participants, to account for potential missing data.

Inclusion criteria comprised an age of 18-65 years, fluency in English and absence of psychiatric or neurological disorders including drug and alcohol addiction or abuse. Participants were screened for these criteria via single items on a self-report questionnaire. To control for potential differences in gaze behavior based on the sex of the interaction partner, only heterosexual participants were included and all participants were interviewed by an experimenter of the same sex.

Eye-tracking data for five participants was erroneously not recorded and therefore not available for analysis. Five participants were tested but did not fulfill inclusion criteria and were excluded from analysis (3 due to non-heterosexual orientation, 1 due to a history of alcohol abuse and 1 due to regular drug use). Two participants were excluded due to missing data on the prescreening questionnaire. The final sample entering analysis consisted of N = 62 participants (33 females, 29 males), 19-63 years of age (M = 26.39, SD = 6.84) (see Table S1).

2.2 Procedure

After completing the prescreening questionnaire, participants were invited to a lab appointment. First, participants gave informed consent and were instructed about the experimental procedure. Next, the eye-tracking headset was put on and calibration was performed using the single-marker calibration choreography. Instructions and calibration lasted approximately 10 minutes, after which the eye tracking recording was started with a button press. Timepoints for interview questions were created by pressing buttons with the respective question label. An additional computer-based task was administered during the course of the experiment related to another research question. The order of tasks was counterbalanced to avoid sequence effects. Thus, the beginning of the interview was not constant across participants, but either 10 or 40 minutes after the start of the experiment. After all interview questions had been asked, the recording was terminated with a button press. Average interview duration over all participants was M = 7.03 min, SD = 4.16 min.

When both tasks had been completed, participants were given the AQ[43] in its shortened version, translated to German and abbreviated to 33 items with most discriminatory power[44]. The AQ is a self-report measure designed to screen for autistic traits in the general population. In this study, it was used as a proxy measure of the broader autism phenotype. Mean of AQ was M = 9.34, SD = 4.82 with a range of 1-25 and four participants scoring

above the threshold of 17 points[44] (see Table S1). Additionally, participants were asked to rate how sympathetic the experimenter was perceived (attractiveness). Attractiveness and gender were used as covariates in the main analysis, in order to control for the confounding effect of these variables.

2.3 Eye-Tracking Interview

A semi-structured interview using a set of 12 questions was conducted during the eye tracking recording. Participant and experimenter were sitting at a table facing each other at a distance of 1m. The participant was wearing an eye-tracking headset described below in more detail. Due to pandemic restrictions, a fully opaque plastic screen was placed between participant and experimenter.

Difficulties in eye contact in ASD and the broader autism phenotype are reported to occur during everyday social interactions[33,45]. Likewise, diagnostic assessment of eye contact is largely based on short interactions between clinician and patient[32]. Thus, similar to previous studies[21,37,46], we created a set of questions meant to resemble everyday conversation such as: "Did you sleep well last night?" or "Did you see a movie last week?". As indicated by findings of reduced eye contact during social interactions[33,45], we expected to observe a reduction of fixations to the eyes for individuals high in autistic traits.

However, we were also interested in whether this effect would be more or less pronounced for certain types of questions. To investigate this, we included both open and closed questions. The interview started out with a number of closed questions (Question 1-6, e.g. "Did you sleep well last night?"), to which participants could answer with a simple "yes" or "no" and continued with a set of open questions, which required a more elaborate answer and possibly more social engagement (Question 7-12, e.g. "What was the most interesting thing that happened to you in the last week?"). We hypothesized that open questions would show a stronger effect of autistic traits on fixations to the eyes compared to closed questions.

Question markers were created using the annotation feature of the pupil labs capture software by pressing a set of keys corresponding to the respective question number. The list of interview questions can be found in the supplementary material and on this papers OSF repository (https://osf.io/y9t5q/).

2.4 Eye-Tracking Equipment

Eye-tracking data was collected using the Pupil Labs Core eye-tracking platform (consisting of a wearable eye-tracking headset and open-source recording and analysis software)[47]. The manufacturer reports the headset to achieve 0.6° visual angle (VA) accuracy under ideal conditions (defined as the angular offset between estimated fixation and corresponding fixation target location). The headset consists of two pupil cameras for gaze estimation and one world camera on the forehead of the participant, recording the participants' field of view.

Eye movements were recorded with 120 Hz sampling frequency. Pupil locations were estimated using the "dark pupil" detection method, which first identifies dark parts and contrasting edges of a converted-to-grayscale image to which an ellipse is fitted[47,48]. We used the 3D pupil detection method, based on a 3D model of the eyes, due to its ability to compensate for movement of the head and slippage of the headset. To further improve data quality, we instructed participants to avoid large movements of the head. To calibrate the headset, we used the single marker calibration choreography with a printed physical marker. The marker was held up in front of the experimenter's face during calibration to ensure appropriate accuracy for this depth level. Mean accuracy over all recordings was $M = 1.70°$ VA, $SD = 0.59°$ VA. At a distance of 1m, this corresponds to an area of 2.96 cm, roughly the diameter of a large coin (i.e. a 2€ coin).

As the experimenter moved during the interview, his face was not a constant size in the participant's visual field. We used the following formula to calculate the degrees of visual angle of the face in the participant's field of view:

$$\alpha = 2 \cdot \arctan\left(\frac{\text{Object size}}{2 \cdot \text{Distance}}\right) \quad (1)$$

Assuming the face to be approximately 20cm in height and 10cm in width, the face took up $\alpha = 11.42°$ VA vertically and $\alpha = 5.72°$ VA horizontally at 1m distance. In the world camera (1280x720 resolution), the face took up approximately 150 pixels vertically and 90 pixels horizontally.

2.5 Data Preprocessing and Face Detection

Fixations were extracted from gaze data using the built-in fixation detector plugin of the pupil labs software. Maximum dispersion was set to 1.5° with a minimum duration of 80 ms and a maximum duration of 220 ms. Blinks (short drops of confidence in pupil detection) were removed using the built-in blink detector plugin. Furthermore, all detected fixations below a threshold of 0.8 confidence were filtered during a separate preprocessing step.

We analyzed the location of fixations on the participant's field of view frame by frame using openCV (*cv2* package) and the *facenet-pytorch* package in python[49]. This package uses a multi-task convolutional neural network (MTCNN) to detect faces in a given image and takes inspiration from David Sandberg's TensorFlow implementation of the FaceNet face recognition model[50,51]. It also includes face detection models pretrained on the VGGFace2 and CASIA-Webface datasets[52,53]. The code to analyze recordings has been published as a GitHub-Repository and is linked to this papers OSF-page (https://osf.io/y9t5q/) or can be accessed on GitHub directly (https://github.com/raimund-buehler/face_detection_small).

Each frame of the world video was first analyzed by the face detection algorithm, yielding the location of the experimenter's face as well as the position of facial landmarks, such as eyes and corners of the mouth. Ellipses were automatically fitted to the face, eyes, and mouth (see Figure 1). The ellipse for the face was fitted using the center, width and height values of the bounding box of the face returned by the face detection algorithm (Figure 1, in yellow). For eyes and mouth (Figure 1, in red and blue), the midpoint between left and right eye or left and right corner of the mouth was computed and used as center for the ellipse. The width value of the facial bounding box was used to fit the ellipse while the height value was scaled down by a factor of 0.3. For eyes and mouth, ellipses were also angled according to the angle formed by the left and right eye or corners of the mouth. Thus, ellipses for eyes and mouth automatically adjusted to the position and size of the experimenter's face in the participant's visual field.

Lastly, if a fixation for a frame was present, it was tested whether it fell onto one of the ellipses (with all ellipses being mutually exclusive). Fixations for each region were then counted and averaged over different time intervals (e.g. total interview duration or question duration), leading to a measure of average fixation rate in Hz (fixations / second) on the respective region during that time interval. This measure served as the dependent variable for further analysis.

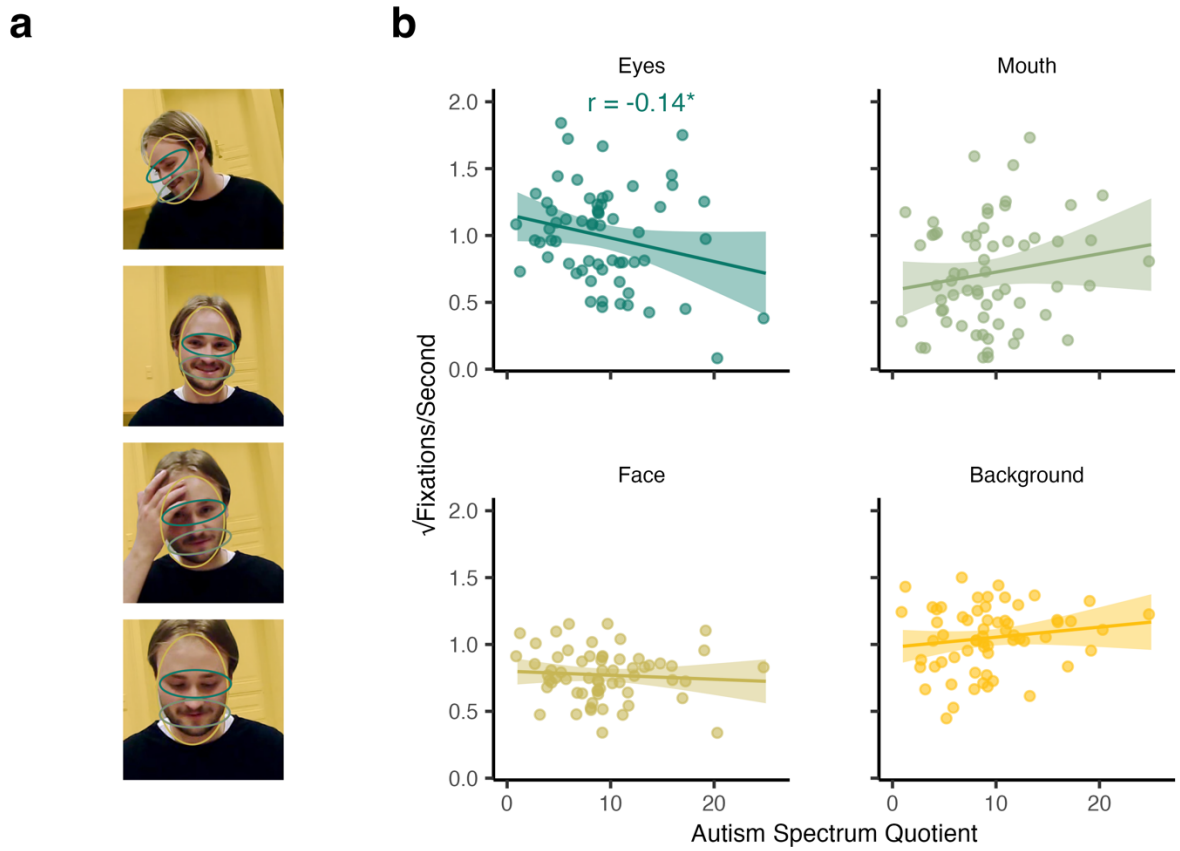

Fig. 1 – a Exemplary output frames of the ellipses fitted to the face and facial regions of the experimenter using cv2 and the facenet-pytorch package. The face and facial landmarks are accurately detected in a range of conditions, for example when looking to the side, in frontal view, when the face is partially occluded or when looking down. **b** Scatterplots for fixation rates on each AOI, averaged over the entire interview. Every data point represents a participant, slightly jittered for improved visibility. Regression lines are depicted as coloured lines and confidence intervals in shaded colors. Fixation rates were square-root transformed to account for non-normality of residuals in linear models. Summary statistics for untransformed fixation rates can be found in the supplementary material.

2.6 Statistical Analysis

The main analyses for this study were preregistered on OSF (https://osf.io/tgcqm). We fit a linear model with fixation rate as dependent variable, an interaction of AOI and AQ as independent variables and gender and perceived attractiveness of the experimenter (Attr.) as covariates. AQ and AOI were defined as interactive effects in order to investigate the effect of AQ at each level of AOI, i.e. specifically for the eye or mouth region. Due to non-normality of residuals (see supplementary results), we used a square root transformation for the dependent variable, as described before for similar eye-tracking data[54]. This transformation was also preregistered.

A number of additional exploratory analyses were conducted, which were not part of the preregistration: First, we investigated the fixation rates on eyes and mouth over the timecourse of the interaction by creating longitudinal plots for each participant. Second, we looked at differences in fixation rates to the eyes when averaged within questions. Lastly, we compared sensitivity to autistic traits for open vs. closed questions by fitting a linear mixed model with AQ and question type (open vs. closed) as fixed effects, including an interaction term between AQ and question type as well as random intercepts by subject.

For fitting linear models, we used the base *lm()* command in R. For mixed models, the *lmer()* command of the *lmerTest*-package[55] was used. The *anova()* command was used to obtain *F*-tests for main effects and

interactions. Significant interactions were further investigated in a simple slope analysis using the *emmeans*-package[56]. Plots were created using *ggplot*[57] and the *interactions*-package[58].

Model assumptions were investigated using plots of residuals versus fitted values, conditional boxplots within levels of AOI, and qqPlots (see supplementary results). Importantly, for a model without transformation of the dependent variable, histograms and qqPlots indicated a significant deviation of the residuals from a normal distribution, confirming the necessity of a square-root transformation for robust results.

We have also created a synthetic version of our dataset using the *synthpop*-package in R. This dataset is a scrambled and anonymized version of the original dataset, which preserves the most important summary aspects[59]. The dataset is freely available on the OSF repository (https://osf.io/y9t5q/) along with a summary of the R code used for plotting, analysis, and model diagnostics. Although the synthetic dataset does not fully capture the results presented here, it permits researchers to reproduce the most important aspects of our analysis.

3 Results

3.1 Main Analysis

Fitting the model described above yielded the following results: the main effect of AOI was significant, $F(3, 238) = 17.37$, $p < .001$, suggesting overall differences in the mean amount spent fixating on the different AOIs. Averaged over the entire interview, participants mostly fixated on the eyes (M = 1.15 fix/s, SD = 0.74) and the background (M = 1.14 fix/s, SD = 0.47) and less on the mouth (M = 0.67 fix/s, SD = 0.65) or the rest of the face (M = 0.63 fix/s, SD = 0.29) (see Table S2). Post-hoc contrasts revealed that significantly more fixations fell on the eye and background region, when compared to the mouth and the rest of face (all $p < 0.001$). However, there were no significant differences between eyes and background or between the mouth and rest of face (all $p > 0.05$). AQ was not significant as a main effect ($F(1, 238) = 0.001$, $p = .97$), but the interaction of AOI × AQ was significant, $F(3, 238) = 2.81$, $p = .04$. This suggests the effect of AQ to be dependent on AOI.

Post-hoc analyses of the interaction revealed it to be driven by a negative effect of AQ for fixations to the eye region (See Figure 2). In a simple slope analysis, this effect was significantly different from zero ($t = -2.17$, $p = .031$). Converting the t-value of the simple slope analysis to a partial correlation coefficient resulted in an effect size of $r = -0.14$. Submitting this effect size to a post-hoc power analysis revealed an actual power of $\beta = 0.58$.

The simple slope analysis also showed that the effect of AQ did not differ significantly from zero for any other AOI (all $p > 0.05$). The main effects of gender ($F(1, 238) = 1.70$, $p = .19$) and attractiveness ($F(1, 238) = 0.06$, $p = .81$) were not significant. Comprehensive results of these analyses in table format can be found in the supplementary material (S3 and S4)

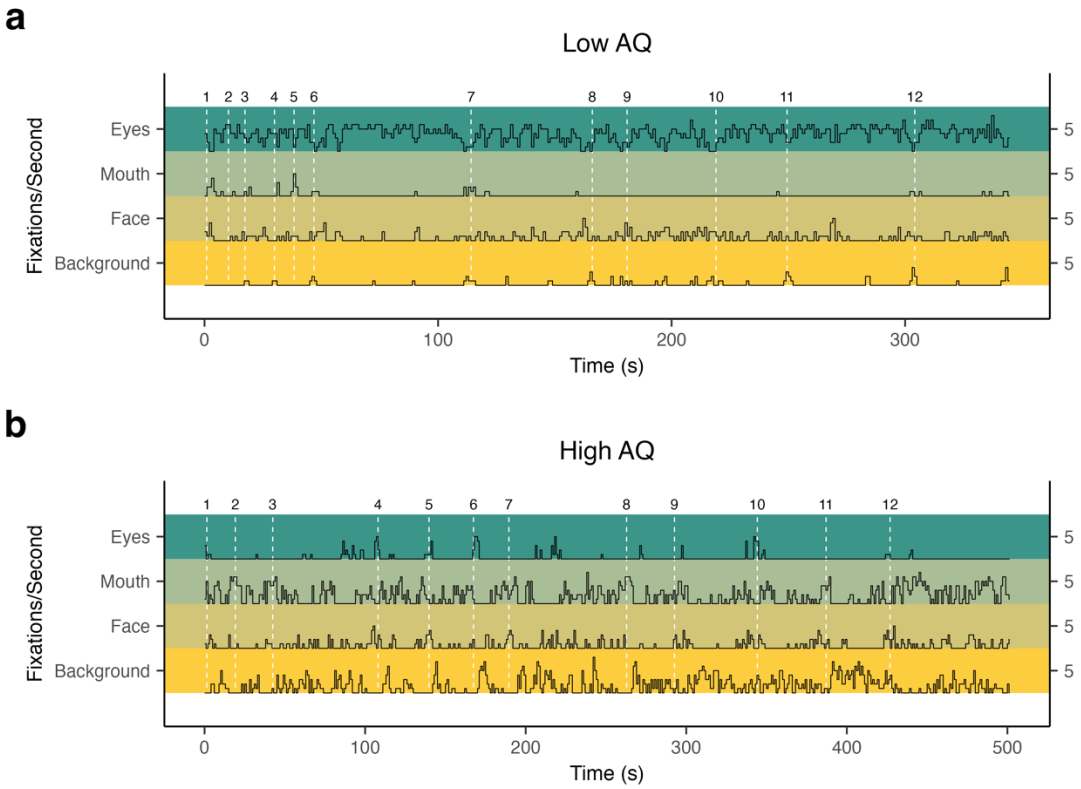

Figure 2 - Longitudinal plots of fixation rates (binned over 1s intervals) for two exemplary participants with low AQ (5 points) and high AQ (17 points). Questions are color coded with vertical lines representing question onset. The high AQ participant is fixating more on the mouth region over the entire interview, whereas the low AQ participant is fixating mostly on the eyes.

3.2 Exploratory Analyses

We conducted several additional exploratory analyses to investigate the timecourse of gaze behavior during the interview. First, we created longitudinal plots for each participant, depicting dynamic fixation rate changes over time and during question intervals. Exemplary plots for two participants with low and high AQ-Score are included in this paper (see Figure 3). As apparent, the high AQ participant fixated mostly on the mouth and less on the eyes over the entire interaction. Conversely, the low AQ participant fixated mostly on the eyes and less on the mouth.

Additionally, we were interested in whether the association of autistic traits and fixation rates to the eyes was a constant effect over the entire interview, or whether it was particularly pronounced for certain parts of the interaction and/or for specific questions (open vs. closed questions).

First, to explore overall differences between questions, we computed the average fixation rate to the eyes for each question and fit a linear mixed model with question number as a main effect and random intercepts by subject. As can be seen in Figure 4, most fixations to the eyes were made during question 4 (“Have you been out of the country in the past week?”) and least for question 9 (“What was the most interesting thing that happened to you in the past week?”). Question 4 (a closed question), differed significantly from question 9, 10 and 11 (open questions, all $p < 0.0001$), with less fixations to the eyes made for open questions.

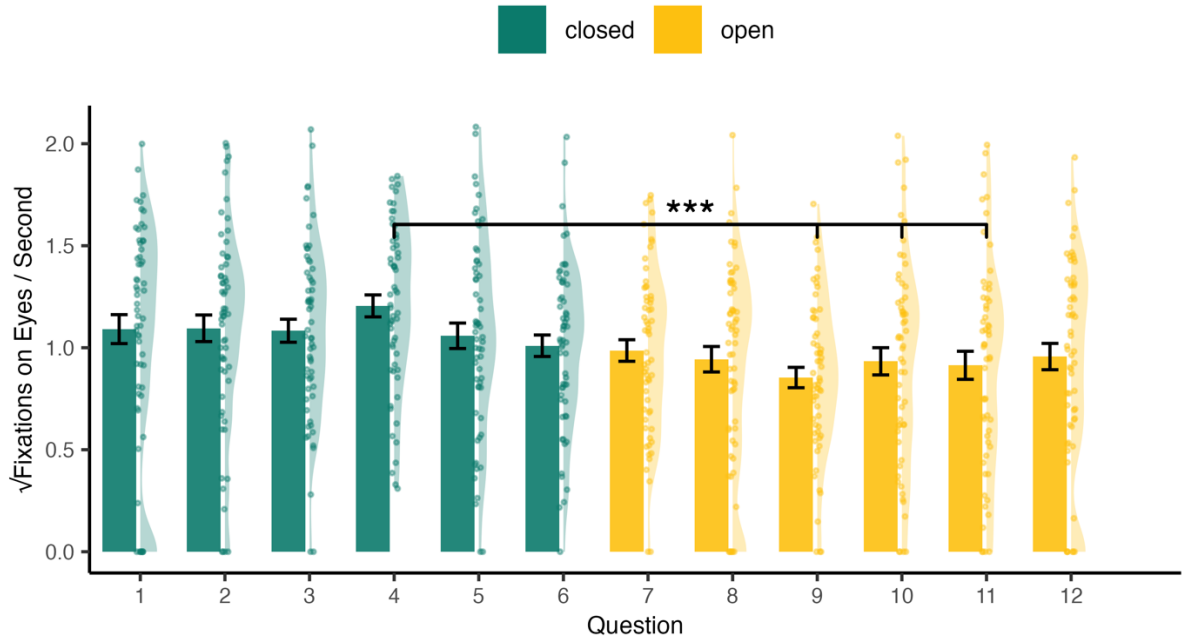

Figure 3 - Average fixation rates within questions for fixations to eyes only. Questions are color coded. Means and standard errors are depicted in red, significant differences are marked with stars ($p < 0.0001$). Fixation rates were square-root transformed to account for non-normality of residuals in linear models.

Lastly, we were also interested in whether open vs. closed questions were differently sensitive to autistic traits. Figure 5 depicts the average fixation rate to the eyes by AQ Score for each question separately. We fit a linear mixed model with question type (open vs. closed) and AQ as main effects, as well as an interaction of question type by AQ and random intercepts by subject (Table S5). As in the model above, the main effect of question type was significant, $F(1, 653.67) = 6.41$, $p = 0.01$, suggesting differences in the average fixation rate on the eyes between open and closed questions with less fixations to the eyes made for open questions. AQ was not a significant main effect in this model, $F(1, 60.62) = 2.12$, $p = 0.15$. Furthermore, the interaction between AQ and question type was not significant $F(1, 658.22) = 0.51$, $p = 0.48$, indicating that open or closed questions were not differentially sensitive to the effect of autistic traits. Thus, while an effect of autistic traits was present when averaging fixation rates over the entire interaction, this was not related to any specific event during the interview, but rather acted as a general effect in this experiment.

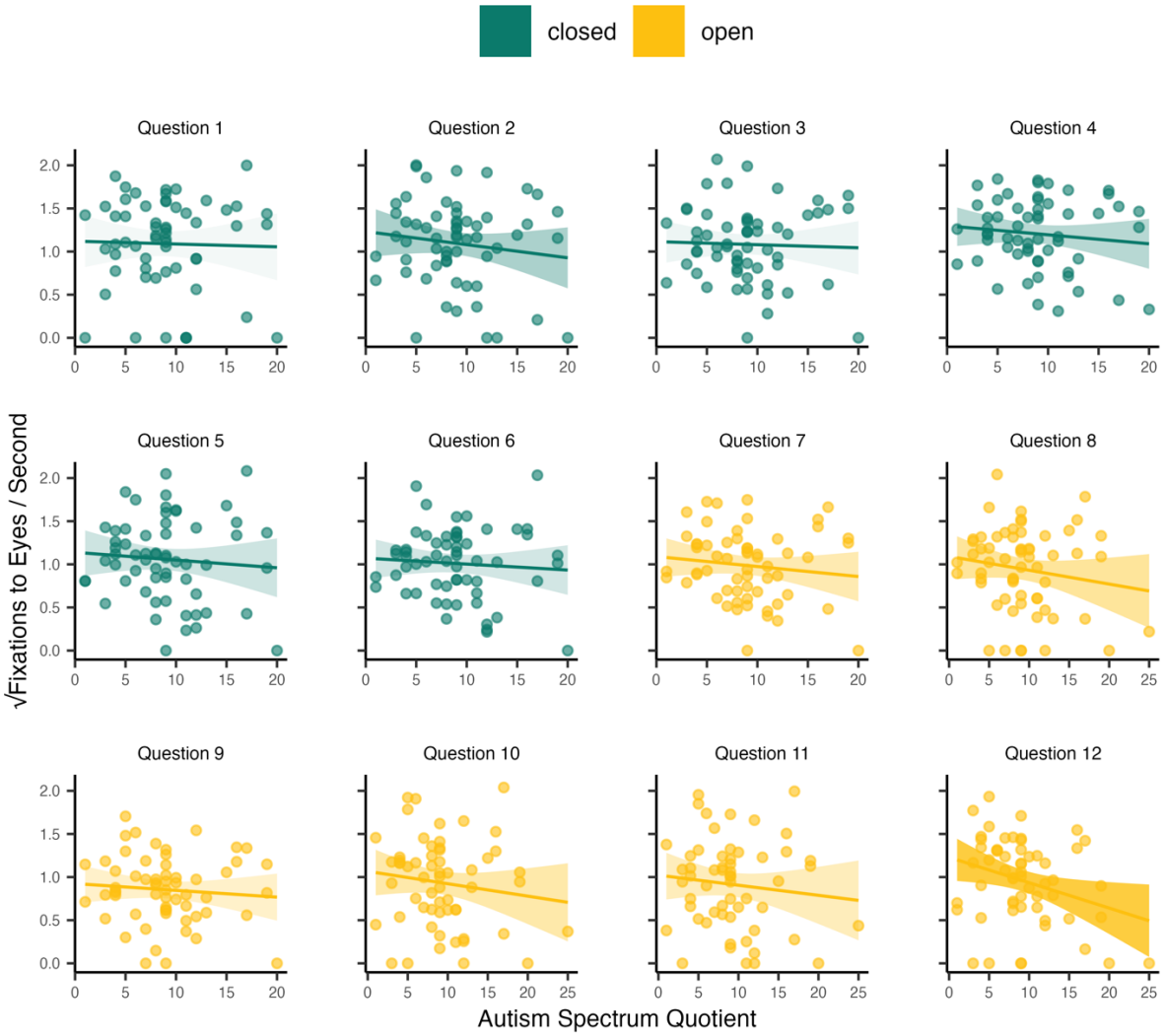

Figure 5 - Fixation rates to the eyes by AQ Score for each question. Question 1 to 6 are closed questions, Question 7 to 12 are open questions. Regression lines are depicted in red. Fixation rates were square-root transformed to account for non-normality of residuals in linear models.

4 Discussion

In light of the need for experiments investigating social attention "in the wild"[31,34,35], the goal of this study was to explore the use of face detection algorithms in a naturalistic experiment, to investigate the effects of autistic traits and to examine the timecourse of gaze behavior over a face-to-face interaction. Using the MTCNN face detection algorithm, we observed a significant negative relationship of AQ on the average fixation rate to the eye region over the entire interview. The effect of autistic traits was not significant for any of the other AOIs. Thus, contrary to some screen-based experiments[7], we did not observe a significant increase in fixations to the mouth for high AQ participants. Some open questions elicited less fixations to the eyes compared to closed questions, potentially because they required a more elaborate response and participants would fixate more on the background while deliberating their reply. However, open questions were not more sensitive towards autistic traits compared to closed questions, as indicated by a non-significant interaction of AQ and question type. Thus, in this study, the reduction of fixations to the eyes was a constant effect over the entire interaction and was not related to the onset of a particular type of question. Generally, the reduction of fixations to the eyes for high AQ participants observed in our study is in line with findings gained in screen-based settings[2,5,10,60] and studies in naturalistic environments, using non-mobile eye tracking or different approaches to AOI coding[14,30,40]. However, our results are in contrast to some studies, which found effects in screen-based, but not in live interaction conditions, for ASD[46,61] or autistic traits[14]. This can likely be attributed to differences in task design and experimental procedure, but more research is needed to determine the precise circumstances in which effects of ASD or autistic traits can be observed.

ASD and other disorders of social behavior are inherently marked by deficits in social interactions in the real world[32,33,45]. It is therefore important to study these behaviors in situations in which they actually occur, involving the presence of another person and including the complexity and richness of real social interactions. Furthermore, one of the major advantages of naturalistic over screen-based experiments is the possibility to study temporal and reciprocal aspects of social attention in everyday interactions. Here, we investigated the timecourse of gaze behavior by plotting fixation rates on AOI's over time and by modelling fixation rates within interview questions. In our study, fixation rates for participants with high AQ were not particularly affected by specific question types but rather reduced over the entire interaction. However, longitudinal analyses of gaze behavior could be expanded to capture interactive aspects more fully, i.e. in response to episodes of speech or the display of emotional expressions by the experimenter.

Emotion recognition algorithms would be a relatively straightforward addition to the face detection algorithm and would allow to automatically classify time intervals when emotional expressions are displayed (smiles, frowns etc.). Episodes of speech could be decoded from an audio signal of the recording[62], although speaker detection is still a challenging aspect. Temporal profiles of gaze behavior, for instance via growth-curve modeling[63], could provide a fruitful statistical approach for longitudinal analysis of gaze behavior. Other studies have also used generalized additive mixed models (GAMM's) to study the temporal distribution of fixations on facial regions[64]. Growth-curve modeling or GAMM's could be used to model the timecourse of gaze behavior in the first seconds after the onset of emotional expressions or speech and investigate whether gaze behavior in these critical time periods differs for ASD or high AQ individuals.

Taken together, our results demonstrate how the utilization of mobile eye-tracking in combination with face detection algorithms offer a promising venue for research in social attention. In light of differences between

findings gained in screen-based settings and naturalistic interactions[65], with some results not generalizing to naturalistic interactions[46,61], developing paradigms that involve realistic social interactions is necessary to generate findings that extend beyond lab settings[66]. Adopting technological advances in face detection algorithms can aid in removing barriers towards such an approach.

Some limitations of the experiment also need to be taken into account. First, our sample did not include participants with an ASD diagnosis, but consisted of a sample of healthy individuals. We used scores on the AQ to capture the broader autism phenotype in the general population. We believe that we have achieved an appropriate range in AQ scores with some participants scoring above the threshold indicative of ASD. Still, results may differ compared to individuals with an ASD diagnosis. Additionally, the study was not at ideal power level ($\beta = 0.58$), as some participants due to exclusion criteria, missing data on questionnaires or missing eye tracking data. However, we believe power and sample size adequate and comparable to other studies in the field.

Furthermore, as participants tend to move their head during conversation (i.e. when nodding) slippage of the headset is to be expected and deteriorates the accuracy of gaze estimation. We have tried to tackle this problem by using the 3D-estimation method, which is able to compensate for movement and headset slippage. We also instructed participants not to move their head too much. Instances of movement should primarily influence the confidence of gaze estimation and relatedly the confidence of fixation detections. Fixations below a certain confidence level were removed during preprocessing. Thus, any potential impact of movement would likely manifest as a general reduction in fixation rates across all AOIs. Given that our findings show an effect of autistic traits specifically for the eye region but not for other AOIs, it is unlikely that movement confounded this result. Furthermore, no significant main effect of AQ was found on overall fixation rates. Nevertheless, the effect of head movement on data quality is important and future studies should address this point more exhaustively.

Third, and relatedly, the size of ellipses fitted to the face, eyes, and mouth was based on the height and width of the facial bounding box returned by the face detection algorithm. Furthermore, while ellipses adjust to changes in the experimenter's head pose, they did not adjust in size. AOI size affects the classification of fixations on AOI's: We used comparatively large AOI's, which compensate for false negative classifications in situations of low accuracy, but at the cost of increasing false positives[67]. Choosing smaller AOI's may have resulted in less fixations classified as falling on the eyes and mouth and more fixations falling on the rest of the face and background, which potentially could have influenced results.

Lastly, the presence of an eye-tracking headset and explicit behavioral instructions may impact the realism of the social situation and, therefore, influence gaze behavior[46,68]. The tradeoff between the invasiveness of the method and the realism of the experimental setup has been described in previous research[69]. While our setup allows for natural interaction, the eye-tracking headset remains noticeable to the participant. More recent approaches, such as appearance-based gaze estimation using OpenFace, which estimate gaze direction based on facial features and head-poses in video recordings, offer promising less invasive alternatives[70]. However, they currently lack the precision of dedicated eye-tracking hardware. Future developments may improve the accuracy of these approaches, offering more ecologically valid methods without compromising data quality.

Despite these limitations, we are confident that the paradigm presented here is sensitive towards interindividual differences in autistic traits and captures important aspects of social attention in face-to-face interactions. Future research may investigate the timecourse of gaze behavior more comprehensively in a longitudinal way and uncover characteristics of gaze in response to certain events during the interaction, i.e. the

display of emotions. We encourage researchers to use or adapt our face detection code or experimental design to achieve these goals.

5 Acknowledgements

We would like to thank the CoBeNe doctoral school and the University of Vienna for providing funds for this project, Daniel Gugerell for providing the eye-tracking equipment, our master students who assisted in data collection (Libor Potočar, Nina Sireč and Serap Özlü), the CogSciHub-SPP team (David Cserjan) and all the participants who took part in the study.

6 Conflict of Interest

The authors declare no conflicts of interest.

7 Data Availability Statement

An anonymized and scrambled version of the dataset as well as a summary of the code used for analysis can be found under the following link (Scientific_Reports/R Analysis): https://osf.io/y9t5q/files/osfstorage. While it does not produce identical results to the ones in the paper, it permits researchers to understand and reconstruct our main analysis.

References

1. Emery, N. J. The eyes have it: the neuroethology, function and evolution of social gaze. *Neurosci. Biobehav. Rev.* **24**, 581–604 (2000).
2. Chita-Tegmark, M. Social attention in ASD: A review and meta-analysis of eye-tracking studies. *Res. Dev. Disabil.* **48**, 79–93 (2016).
3. Dawson, G., Webb, S. J. & McPartland, J. Understanding the nature of face processing impairment in autism: insights from behavioral and electrophysiological studies. *Dev. Neuropsychol.* **27**, 403–424 (2005).
4. Schultz, R. T. Developmental deficits in social perception in autism: the role of the amygdala and fusiform face area. *Int. J. Dev. Neurosci.* **23**, 125–141 (2005).
5. Frazier, T. W. *et al.* A Meta-Analysis of Gaze Differences to Social and Nonsocial Information Between Individuals With and Without Autism. *J. Am. Acad. Child Adolesc. Psychiatry* **56**, 546–555 (2017).
6. Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G. & Jones, W. Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature* **459**, 257–261 (2009).
7. Klin, A., Jones, W., Schultz, R., Volkmar, F. & Cohen, D. Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Arch. Gen. Psychiatry* **59**, 809–816 (2002).
8. Pelphrey, K. A. *et al.* Visual Scanning of Faces in Autism. *J. Autism Dev. Disord.* **32**, 249–261 (2002).
9. Jones, W., Carr, K. & Klin, A. Absence of Preferential Looking to the Eyes of Approaching Adults Predicts Level of Social Disability in 2-Year-Old Toddlers With Autism Spectrum Disorder. *Arch. Gen. Psychiatry* **65**, 946–954 (2008).
10. Guillon, Q., Hadjikhani, N., Baduel, S. & Rogé, B. Visual social attention in autism spectrum disorder: Insights from eye tracking studies. *Neurosci. Biobehav. Rev.* **42**, 279–297 (2014).
11. Hessels, R. S., Holleman, G. A., Cornelissen, T. H. W., Hooge, I. T. C. & Kemner, C. Eye contact takes two – autistic and social anxiety traits predict gaze behavior in dyadic interaction. *J. Exp. Psychopathol.* **9**, (2018).
12. Freeth, M., Foulsham, T. & Kingstone, A. What Affects Social Attention? Social Presence, Eye Contact and Autistic Traits. *PLOS ONE* **8**, e53286 (2013).
13. Nummenmaa, L., Engell, A. D., Von dem Hagen, E., Henson, R. N. A. & Calder, A. J. Autism spectrum traits predict the neural response to eye gaze in typical individuals. *NeuroImage* **59**, 3356–3363 (2012).
14. Vabalas, A. & Freeth, M. Brief report: Patterns of eye movements in face to face conversation are associated with autistic traits: Evidence from a student sample. *J. Autism Dev. Disord.* **46**, 305–314 (2016).
15. Wegner-Clemens, K., Rennig, J. & Beauchamp, M. S. A relationship between Autism-Spectrum Quotient and face viewing behavior in 98 participants. *PLOS ONE* **15**, e0230866 (2020).
16. Laidlaw, K. E. W., Foulsham, T., Kuhn, G. & Kingstone, A. Potential social interactions are important to social attention. *Proc. Natl. Acad. Sci.* **108**, 5548–5553 (2011).
17. Kleberg, J. L. *et al.* Autistic Traits and Symptoms of Social Anxiety are Differentially Related to Attention to Others' Eyes in Social Anxiety Disorder. *J. Autism Dev. Disord.* **47**, 3814–3821 (2017).
18. Nayar, K., Shic, F., Winston, M. & Losh, M. A constellation of eye-tracking measures reveals social attention differences in ASD and the broad autism phenotype. *Mol. Autism* **13**, 18 (2022).
19. Risko, E. F., Laidlaw, K. E. W., Freeth, M., Foulsham, T. & Kingstone, A. Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. *Front. Hum. Neurosci.* **6**, 143 (2012).
20. Risko, E. F., Laidlaw, K. E. W., Freeth, M., Foulsham, T. & Kingstone, A. Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. *Front. Hum. Neurosci.* **6**, 143 (2012).
21. Freeth, M., Foulsham, T. & Kingstone, A. What Affects Social Attention? Social Presence, Eye Contact and Autistic Traits. *PLOS ONE* **8**, e53286 (2013).
22. Chevallier, C. *et al.* Measuring social attention and motivation in autism spectrum disorder using eye-tracking: Stimulus type matters. *Autism Res.* **8**, 620–628 (2015).
23. Frischen, A., Bayliss, A. P. & Tipper, S. P. Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychol. Bull.* **133**, 694–724 (2007).
24. Hernandez, N. *et al.* Exploration of core features of a human face by healthy and autistic adults analyzed by visual scanning. *NEUROPSYCHOLOGIA* **47**, 1004–1012 (2009).
25. Leppänen, J. M. & Nelson, C. A. Tuning the developing brain to social signals of emotions. *Nat. Rev. Neurosci.* **10**, 37–47 (2009).
26. Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y. & Plumb, I. The "Reading the Mind in the Eyes" Test Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism. *J. Child Psychol. Psychiatry* **42**, S0021963001006643 (2001).
27. Itier, R. J., Villate, C. & Ryan, J. D. Eyes always attract attention but gaze orienting is task-dependent: Evidence from eye movement monitoring. *Neuropsychologia* **45**, 1019–1028 (2007).
28. Birmingham, E., Bischof, W. F. & Kingstone, A. Social Attention and Real-World Scenes: The Roles of Action, Competition and Social Content. *Q. J. Exp. Psychol.* **61**, 986–998 (2008).
29. Foulsham, T., Cheng, J. T., Tracy, J. L., Henrich, J. & Kingstone, A. Gaze allocation in a dynamic situation: Effects of social status and speaking. *Cognition* **117**, 319–331 (2010).
30. Hessels, R. S., Holleman, G. A., Cornelissen, T. H. W., Hooge, I. T. C. & Kemner, C. Eye contact takes two – autistic and social anxiety traits predict gaze behavior in dyadic interaction. *J. Exp. Psychopathol.* **9**, (2018).
31. Risko, E. F., Richardson, D. C. & Kingstone, A. Breaking the fourth wall of cognitive science: Real-world social attention and the dual function of gaze. *Curr. Dir. Psychol. Sci.* **25**, 70–74 (2016).
32. MacDonald, K. Patient-Clinician Eye Contact: Social Neuroscience and Art of Clinical Engagement. *Postgrad. Med.* **121**, 136–144 (2009).
33. Trevisan, D. A., Roberts, N., Lin, C. & Birmingham, E. How do adults and teens with self-declared Autism Spectrum Disorder experience eye contact? A qualitative analysis of first-hand accounts. *PLOS ONE* **12**, e0188446 (2017).
34. Schilbach, L. *et al.* Toward a second-person neuroscience. *Behav. Brain Sci.* **36**, 393–414 (2013).
35. Schilbach, L. Eye to eye, face to face and brain to brain: novel approaches to study the behavioral dynamics and neural mechanisms of social interactions. *Curr. Opin. Behav. Sci.* **3**, 130–135 (2015).
36. Jongerius, C. *et al.* Eye-tracking glasses in face-to-face interactions: Manual versus automated assessment of areas-of-interest. *Behav. Res. Methods* **53**, 2037–2048 (2021).
37. Auyeung, B. *et al.* Oxytocin increases eye contact during a real-time, naturalistic social interaction in males with and without autism. *Transl. Psychiatry* **5**, (2015).

38. Chong, E. *et al.* Detecting Gaze Towards Eyes in Natural Social Interactions and Its Use in Child Assessment. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **1**, 1–20 (2017).
39. de Belen, R. A. J., Bednarz, T., Sowmya, A. & Del Favero, D. Computer vision in autism spectrum disorder research: a systematic review of published studies from 2009 to 2019. *Transl. Psychiatry* **10**, (2020).
40. Falck-Ytter, T. Gaze performance during face-to-face communication: A live eye tracking study of typical children and children with autism. *Res. Autism Spectr. Disord.* **17**, 78–85 (2015).
41. Noris, B., Nadel, J., Barker, M., Hadjikhani, N. & Billard, A. Investigating gaze of children with ASD in naturalistic settings. (2012).
42. Hanley, M. *et al.* The use of eye-tracking to explore social difficulties in cognitively able students with autism spectrum disorder: A pilot investigation. *Autism* **19**, 868–873 (2015).
43. Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. & Clubley, E. The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Malesand Females, Scientists and Mathematicians. *J. Autism Dev. Disord.* **31**, 5–17 (2001).
44. Freitag, C. M. *et al.* Evaluation der deutschen Version des Autismus-Spektrum-Quotienten (AQ) - die Kurzversion AQ-k. *Z. Für Klin. Psychol. Psychother.* **36**, 280–289 (2007).
45. Senju, A. & Johnson, M. H. Atypical eye contact in autism: Models, mechanisms and development. *Neurosci. Biobehav. Rev.* **33**, 1204–1214 (2009).
46. Cañigueral, R., Ward, J. A. & Hamilton, A. F. de C. Effects of being watched on eye gaze and facial displays of typical and autistic individuals during conversation. *Autism* **25**, 210–226 (2021).
47. Kassner, M., Patera, W. & Bulling, A. Pupil. in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication* 1151–1160 (ACM, New York, NY, USA, 2014). doi:10.1145/2638728.2641695.
48. Świrski, L., Bulling, A. & Dodgson, N. Robust real-time pupil tracking in highly off-axis images. in *Proceedings of the Symposium on Eye Tracking Research and Applications* 173–176 (ACM, New York, NY, USA, 2012). doi:10.1145/2168556.2168585.
49. Tim Esler. facenet-pytorch. *GitHub Repository* https://github.com/timesler/facenet-pytorch (2020).
50. David Sanberg. facenet. https://github.com/davidsandberg/facenet (2016).
51. Schroff, F., Kalenichenko, D. & Philbin, J. FaceNet: A unified embedding for face recognition and clustering. in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 815–823 (IEEE, 2015). doi:10.1109/CVPR.2015.7298682.
52. Cao, Q., Shen, L., Xie, W., Parkhi, O. M. & Zisserman, A. VGGFace2: A dataset for recognising faces across pose and age. (2017).
53. Yi, D., Lei, Z., Liao, S. & Li, S. Z. Learning Face Representation from Scratch. (2014).
54. Auyeung, B. *et al.* Oxytocin increases eye contact during a real-time, naturalistic social interaction in males with and without autism. *Transl. Psychiatry* **5**, (2015).
55. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest Package: Tests in Linear Mixed Effects Models. *J. Stat. Softw.* **82**, (2017).
56. Lenth, R., Singmann, H., Love, J., Buerkner, P. & Herve, M. Package 'emmeans'. (2019).
57. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2016).
58. Long, J. A. interactions: Comprehensive, User-Friendly Toolkit for Probing Interactions. (2019).
59. Nowok, B., Raab, G. M. & Dibben, C. synthpop : Bespoke Creation of Synthetic Data in R. *J. Stat. Softw.* **74**, (2016).
60. Chen, F. S. & Yoon, J. M. D. Brief Report: Broader Autism Phenotype Predicts Spontaneous Reciprocity of Direct Gaze. *J. Autism Dev. Disord.* **41**, 1131–1134 (2011).
61. Grossman, R. B., Zane, E., Mertens, J. & Mitchell, T. Facetime vs. Screentime: Gaze Patterns to Live and Video Social Stimuli in Adolescents with ASD. *Sci. Rep.* **9**, 12643 (2019).
62. Rogers, S. L., Speelman, C. P., Guidetti, O. & Longmuir, M. Using dual eye tracking to uncover personal gaze patterns during social interaction. *Sci. Rep.* **8**, 1–9 (2018).
63. Del Bianco, T. *et al.* Temporal Profiles of Social Attention Are Different Across Development in Autistic and Neurotypical People. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **6**, 813–824 (2021).
64. Maes, P., Stercq, F. & Kissine, M. Brief report: Temporal distribution of visual attention between the eyes and mouth in young autistic children. *Res. Autism Spectr. Disord.* **110**, 102292 (2024).
65. Laidlaw, K. E. W., Foulsham, T., Kuhn, G. & Kingstone, A. Potential social interactions are important to social attention. *Proc. Natl. Acad. Sci.* **108**, 5548–5553 (2011).
66. Leong, V. & Schilbach, L. The promise of two-person neuroscience for developmental psychiatry: using interaction-based sociometrics to identify disorders of social interaction. *Br. J. Psychiatry* **215**, 636–638 (2019).
67. Vehlen, A., Standard, W. & Domes, G. How to choose the size of facial areas of interest in interactive eye tracking. *PLOS ONE* **17**, e0263594 (2022).
68. Cañigueral, R., Hamilton, A. & Ward, J. A. Don't Look at Me, I'm Wearing an Eyetracker! in Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers 994–998 (ACM, Singapore Singapore, 2018). doi:10.1145/3267305.3274123.
69. Valtakari, N. V. *et al.* Eye tracking in human interaction: Possibilities and limitations. *Behav. Res. Methods* **53**, 1592–1608 (2021).
70. Zhang, X., Sugano, Y., Fritz, M. & Bulling, A. Appearance-based gaze estimation in the wild. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 4511–4520 (2015).

Supplementary Material

Supplementary Tables

Table S1. Participant characteristics

Variable	Mean (SD)	Range
Age	26.39 (6.84)	19.00-63.00
AQ Score	9.34 (4.82)	1.00-25.00
Attract. Rating	5.57 (0.76)	4.00-7.00

	Level	N (%)
Gender	*Male*	29 (46.8)
	Female	33 (53.2)
Vision	*Not Impaired*	43 (69.4)
	Glasses	19 (30.6)
English Level	*C1 - C2*	45 (72.6)
	B1 - B2	14 (22.6)
	Native Speaker	3 (4.8)

Table S2. Summary statistics for the average fixation rate on the four AOI's over all participants

AOI	Mean	SD	Range
Eyes	1.15	0.74	0.01 - 3.39
Mouth	0.67	0.65	0.01 - 3.00
Face	0.63	0.29	0.12 - 1.33
Background	1.14	0.47	0.20 - 2.25

Table S3. F-tests (ANOVA) for the linear model (LM) fitted to the average fixation rate

Predictors	Df	Sum Sq	Mean Sq	F	p
AQ	1	0.0001	0.0001	0.001	0.97
AOI	3	4.94	1.65	17.37	**<0.001**
Gender	1	0.16	0.16	1.70	0.19
Attractiveness	1	0.01	0.01	0.06	0.81
AQ:AOI	3	0.80	0.27	2.81	**0.04**
Residuals	238	22.56	0.09		NA

Note. AQ: *Autism Spectrum Quotient*, AOI: *Area of Interest*, Attr.: *Attractiveness of Interviewer*. Significant p values are marked in bold.

Table S4 – Simple slope analysis for the effect of AQ at different levels of AOI

AOI	AQ Slope	SE	DF	LCL	UCL	t	p
Eyes	-0.018	0.01	238.00	-0.034	-0.002	-2.17	**0.031**
Mouth	0.014	0.01	238.00	-0.003	0.030	1.65	0.100
Face	-0.003	0.01	238.00	-0.019	0.013	-0.39	0.700
Background	0.007	0.01	238.00	-0.009	0.024	0.90	0.368

Table S5 - F-tests (ANOVA) for the linear mixed model (LMM) fitted to the average fixation rate within questions. DenDF values were approximated using Satterthwaite's Method.

	Sum Sq	Mean Sq	NumDF	DenDF	F value	p
AQ	0.17	0.17	1.00	60.62	2.12	0.15
Question Type	0.52	0.52	1.00	653.67	6.41	**0.01**
AQ:Question Type	0.04	0.04	1.00	657.22	0.51	0.48

Supplementary Methods

Oral instructions for participants:
"We are now going to have a short conversation. Please make yourself comfortable, so you will be able to sit in the same position for the next few minutes. Please try to avoid large movements of your head, small movements like nodding are ok. I will ask you a couple of questions and later on, you will have the opportunity to ask me some questions as well. The idea is to have as relaxed and natural a conversation as possible. If you are ready, I will now start with the first question."

1. Did you sleep well last night?
2. Did you dream about anything?
3. Have you eaten at a restaurant in the past week?
4. Have you been out of the country in the past week?
5. Have you read any book in the past week?
6. What was the last movie or TV show that you watched?
7. Could you tell me about some things that you did last weekend?
8. Can you tell me about your plans for the next weekend?
9. What was the most interesting thing that happened to you in the past week?
10. When you woke up this morning, were you looking forward to participating in this experiment or was it more of an obligation for you?
11. Could you tell me a little bit about your journey to the lab?
12. Do you have any questions for me?

Supplementary Results

Model Diagnostics

Model assumptions for the preregistered model (using a square-root transformation of the dependent variable) were checked in the following way: Linearity was verified using plots of residuals versus fitted values which was judged to appear sufficiently random and, thus, indicated no deviation from linearity (see Figure S1).

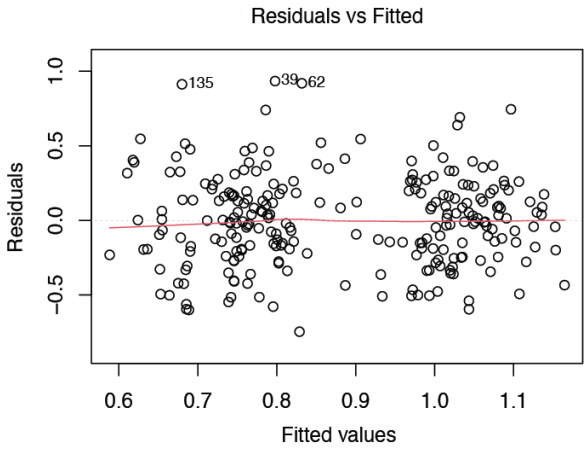

Figure S1 – Residuals versus fitted values for the preregistered linear mixed model described in Section 2.6. Points are clustered as a horizontal band around the zero line. Deviations from linearity would appear as "curved" shapes. Heteroscedasticity could be recognized by points clustering closer to zero for low versus high fitted values (a funnel shape along the x-axis).

Homoscedasticity (homogeneity of variances) was investigated using conditional boxplots (see Figure S2) and a number of homogeneity tests. Variances appear unequal for the different AOIs in boxplots. Levene's, Brown-Forsythe and Kruskal-Wallis tests confirm this assumption (all *p*s < .001). However, the ratio of largest to smallest variance over AOI levels was 6.70, which is in the range of the recommended cutoff for unequal variances across groups[68].

Normality of residuals was investigated using qqPlots, histograms and the Shapiro-Wilk test (W = 0.99, $p = .12$), all of which indicated no deviation from normal distribution (Figure S3). Importantly, when fitting a model without transformation of the dependent variable, both qqPlots and the Shapiro-Wilk test (W = 0.93, $p < .0001$) indicated a significant deviation from normal distribution, corroborating the need of a square-root transformation to account for positive skew.

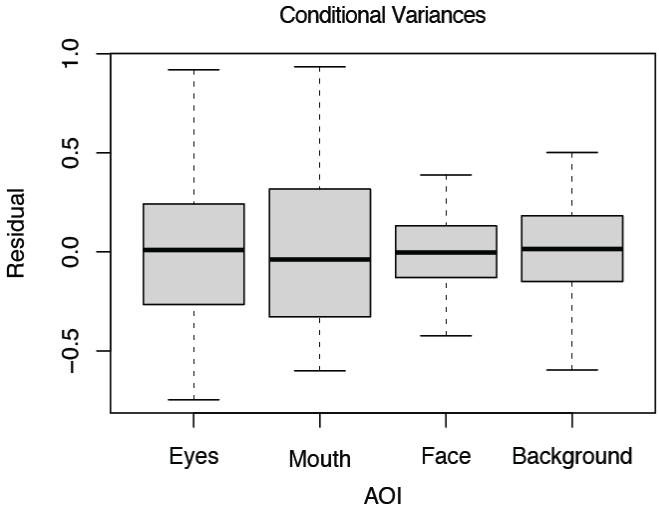

Figure S2 – Conditional boxplots for the variance of residuals in each level of area of interest (AOI). Variances appear slightly unequal for the different levels.

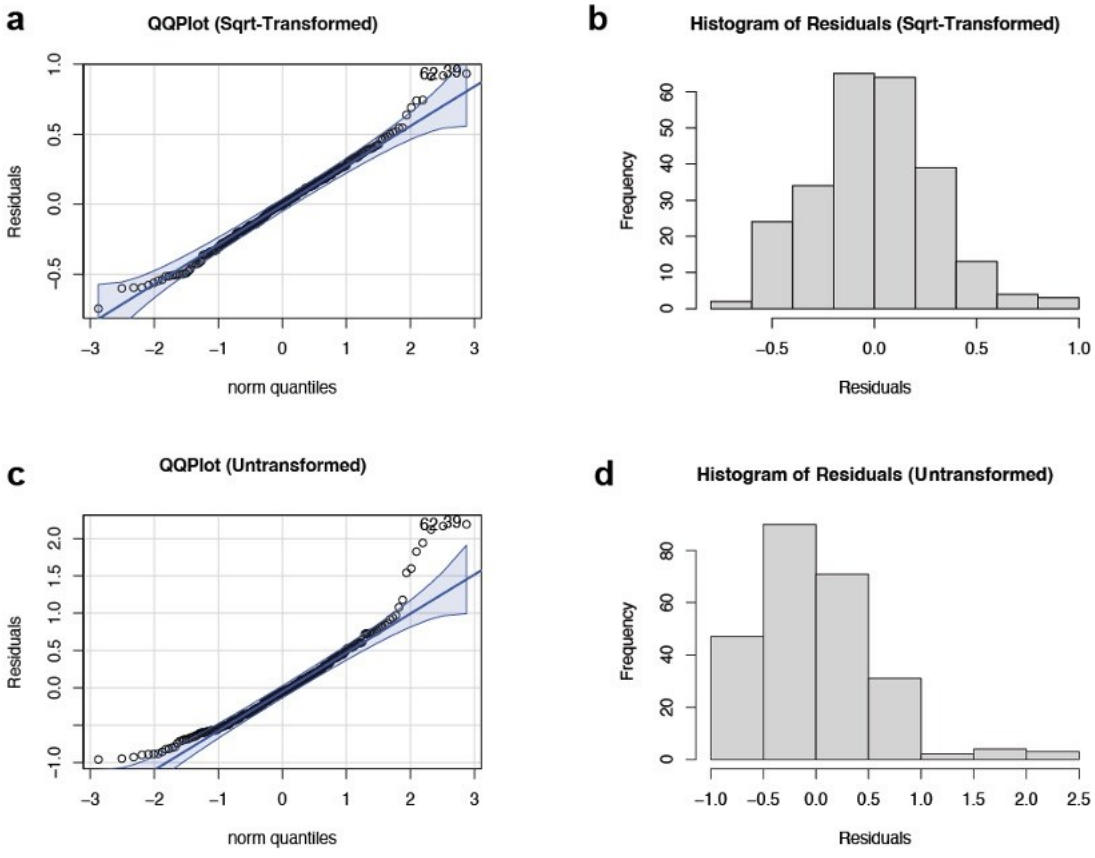

Figure S3 – QQPlots and histograms of the residuals using a square-transformation of the dependent variable (**a** and **b**) and the untransformed model (**c** and **d**). Residuals deviate from the reference line for the untransformed model (**c**) and show positive skew in the histogram (**d**), corroborating the need for a square-root transformation.

Chapter 2 – Autistic Traits relate to reduced Reward Sensitivity in Learning from Point-Light Displays (PLD's)

Raimund Buehler[1], Libor Potocar[2], Nace Mikus[3], Giorgia Silani[1]

[1]Department of Clinical and Health Psychology, Faculty of Psychology, University of Vienna, Austria

[2]PROMENTA Research Center, Department of Psychology, University of Oslo, Oslo, Norway

[3]Interacting Minds Centre, Aarhus University, Denmark

Corresponding authors:

Raimund Buehler
raimund.buehler@univie.ac.at
+43-1-4277-47241
Liebiggasse 5, 1010 Vienna, Austria

Giorgia Silani
giorgia.silani@univie.ac.at
+43-1-4277-47223
Liebiggasse 5, 1010 Vienna, Austria

Abstract

A number of studies have linked autistic traits to difficulties in learning from social (vs. non-social) stimuli. However, these stimuli are often difficult to match on low-level visual properties, which is especially important given the impact of autistic traits on sensory processing. Additionally, studies often fail to account for dissociable aspects of the learning process in the specification of model parameters (learning rates and reward sensitivity). Here, we investigate whether learning deficits in individuals with high autistic traits exhibit deficits when learning from facial point-light displays (PLDs) depicting emotional expressions. Social and non-social stimuli were created from random arrangements of the same number of point-lights and carefully matched on low-level visual properties. Neurotypical participants (N = 63) were assessed using the Autism Spectrum Quotient (AQ) and completed a total of 96 trials in a reinforcement learning task. Although linear multilevel modeling did not indicate learning deficits, preregistered computational modeling using a Rescorla-Wagner framework revealed that higher autistic traits were associated with reduced reward sensitivity in the win domain, demonstrating an attenuated response to received feedback during learning. These findings suggest that autistic traits can significantly impact learning from PLD feedback beyond a general deficit in learning rates.

1 Introduction

Autism spectrum disorder (ASD) represents a wide range of neuro-developmental conditions characterized by deficits in social communication and interaction and repetitive or restricted patterns of behaviors, interests, or activities. In line with the notion of universally accepted heterogeneity of autism (Jeste & Geschwind, 2014; Mottron & Bzdok, 2020), the Diagnostic and Statistical Manual of mental disorders V (DSM-V) currently adopts a dimensional approach to ASD (American Psychiatric Association, 2013), highlighting that ASD represents a spectrum extending into the general population.

While several theories have been put forward to explain ASD symptoms, and relatedly autistic traits in the general population, one of the most promising approaches to its etiology is to describe the condition as a learning deficit (Van de Cruys et al., 2014). For example, predictive processing construes ASD symptoms as a consequence of a core deficit in learning about the contingencies of the environment, often resulting in imprecise predictions. This is especially relevant when stimuli are complex and dynamic, i.e. when dealing with social situations (Van de Cruys et al., 2014). Difficulties in social domains are thus often interpreted to result from a failure to adapt *learning rates* (the weighting of prediction errors) to different contexts, i.e. by overestimating the volatility of the environment and assigning excessive weight to recent information (Lawson et al., 2017). This can in turn result in difficulties with learning and behavior in social interactions, however supposed deficits are not necessarily confined to social situations but rather described as a general learning deficit (Lawson et al., 2017; Palmer et al., 2017).

A related but alternative account is the *slow updating hypothesis*, which suggests that autistic individuals integrate new information more slowly r (Lieder et al., 2019; Vishne et al., 2021). For example, Lieder and colleagues found that autistic participants underweighted recent sensory input in a perceptual discrimination task, while Vishne and colleagues showed reduced trial-by-trial error correction in a sensorimotor synchronization task. Both studies indicate that rather than an increased reliance on prediction errors, autistic learning difficulties may stem from slower short-term adjustments, potentially affecting adaptation in both social and non-social contexts.

A competing account of the learning difficulties in ASD arises from the *social motivation hypothesis* (Chevallier et al., 2012). This theory proposes that ASD can be characterized by a fundamental difference in the *processing of social rewards*, similarly leading to less optimal learning in social contexts, but due to slightly different reasons (Lin et al., 2012; Solomon et al., 2015, for a review see Schuetze et al., 2017). According to this perspective, learning deficits mainly emerge as a consequence of attenuated sensitivity to rewarding social stimuli. Learning difficulties are therefore confined to social contexts, in contrast to the notion of a general learning deficit described above.

Hypotheses derived from both accounts are often tested using reinforcement learning models (Lawson et al., 2017; Palmer et al., 2017; Sutton & Barto, 2018). In these models, choice behavior in response to rewarding feedback is investigated in a trial-by-trial approach. Different parameters are then estimated on the subject level and compared between groups. These parameters permit making conclusions about the underlying causes of learning deficits. For instance, studies often tend to focus on group differences in the learning rate α (Zhang et al., 2020). This parameter captures the weighting of prediction errors and measures the general *speed* at which rewards affect behavior during learning. However, reinforcement learning models can also be specified with additional parameters, such as the reward sensitivity ρ. This parameter measures the *internal value* of a specific type of reward (Huys et al., 2013). Distinguishing these components in the investigation of learning deficits is important,

as they can be used to make inferences about underlying cognitive processes. For example, attempts to disentangle the contribution of each parameter have been made in computational models of depression, where anhedonia was shown to affect reward sensitivity, whereas dopamine manipulation affected learning rates (Huys et al., 2013). Thus, both parameters are theoretically distinguishable and can be used to inform theoretical assumptions. In the context of autistic traits, reduced reward sensitivity in the social condition should reflect lower internal values for social stimuli, in line with theoretical accounts of attenuated social motivation. In contrast, differences in learning rates extending to non-social stimuli would convey a more general learning deficit, which may be particularly relevant in social situations, but not necessarily confined to them.

To test the specificity of learning difficulties in ASD, studies have so far often relied on pictures or videos of emotional expressions, which are then compared to non-social stimuli such as money or objects. While some studies found support for the domain-specific hypothesis of an impairment exclusively for social contexts (Chambon et al., 2017; Ewbank et al., 2017; Lin et al., 2012; Palumbo et al., 2015; von der Lühe et al., 2016), other results suggest deficits also in non-social conditions (Scott-Van Zeeland et al., 2010; Solomon et al., 2015). Importantly, these deficits seem to generalize to autistic traits in the general population as measured by the autism spectrum quotient (AQ) (Baron-Cohen et al., 2001; Freitag et al., 2007), both in regards to a general learning deficit (Coll et al., 2020; Goris et al., 2021; Karvelis et al., 2017) as well as an impairment confined exclusively to social contexts (Bianco et al., 2020; Heerey, 2014; Sevgi et al., 2020).

A major limitation and confounding aspect in these studies concerns the matching of social and non-social stimuli on low-level visual properties, which may contribute to the discrepant results about the specificity of learning deficits. Visual stimuli used for the non-social condition such as objects or shapes (Ewbank et al., 2017; Hurlemann et al., 2010; Kinard et al., 2020) are difficult to match in aspects like color, texture, and luminance. This can potentially impact participants' responses (Freitag et al., 2008), especially in autistic individuals for which atypicalities often arise at the sensory-perceptual level (Mayer, 2017). Therefore, matching stimuli on low level visual properties is critical to properly interpret emerging differences. For this study, we have thus created a novel set of stimuli by drawing on the field of *biological motion*. This line of research has established the use of point-light displays (PLDs), a class of low-level perceptual stimuli consisting of white points on a black background, which offer separation of social and emotional content from other visual features of the stimulus. It has been demonstrated that facial emotional expressions can be recognized from PLDs (Bassili, 1978; Bidet-Ildei et al., 2020; Sowden et al., 2021) and that such stimuli can reliably activate brain regions associated with face processing, such as the fusiform face area (FFA) (Atkinson et al., 2012). Furthermore, neurotypical children exhibit visual preference for PLDs depicting biological motion when compared to autistic children (Kaliukhovich et al., 2021; Klin et al., 2009) and high autistic traits relate to impaired emotion recognition from facial PLDs (Keating et al., 2022). Importantly, some researchers have connected deficits in emotion recognition to social motivation and reward processing (Dawson et al., 2005).

To our knowledge, PLDs have so far not been used to investigate learning from social and non-social stimuli. In this study, we therefore examined differences in learning related to autistic traits using PLDs in a reinforcement learning task. In the social condition, participants were presented with facial PLDs depicting positive or negative emotions (i.e. happiness or anger) to learn an underlying association. In the non-social condition, initially random arrangements of the same number of point-lights formed conventional symbols of success or failure (a checkmark or across). Our goal was to investigate how autistic traits would impact learning

from social and non-social PLDs. Additionally, we were interested in whether variation in parameters of the reinforcement learning model (learning rate and reward sensitivity) in different conditions (non-social vs social) could lend support to competing theoretical accounts of learning deficits in autism (e.g. general learning deficit or attenuated sensitivity to social rewards). We expected participants with high autistic traits to perform worse in the social condition and no differences in the nonsocial condition, but we did not make predictions as to which parameter of the reinforcement learning model precisely was affected.

2 Methods

2.1 Sample

The sample consisted of typically developing (TD) individuals, which were contacted via a participation recruitment platform of the University of Vienna. Participants were required to fulfill the following inclusion criteria: a) age between 18 and 65 years, b) heterosexual orientation, c) proficiency in English, d) no drug or alcohol addiction or regular drug use, e) no psychiatric or neurological condition. As the experiment was part of a larger project, some of the exclusion criteria are related to another task and not directly relevant for this study.

From the total of 74 recruited individuals, three participants were excluded due to missing data for AQ scores and five participants were dropped based on exclusion criteria. Three individuals were excluded from analyses because of missing data for the task, leading to a final sample of N = 63 participants for analysis. Participants received either a financial compensation of 10€ or 4 study credits.

2.2 Measures - AQ

To measure autistic traits, a German shortened version of the Autism Spectrum Quotient (AQ-k; Freitag et al., 2007), widely used in research and clinical practice, was used. AQ-k contains 33 items (e.g. "*I prefer to do things on my own rather than with others.*") and is suitable for adults and adolescents aged 16 years and above with normal intellectual functioning. For screening purposes of ASD in clinical practice, a cut-off of 17 was proposed (Freitag et al., 2007). In the present study, item responses were scored using a binary system, where endorsement of an autistic trait is scored with one point, while the opposite response is scored with zero, resulting in a maximum score of 33 (Ruzich et al., 2015). In the present sample, we report reliability scores as Cronbach's alpha (α) = .83 and McDonald's omega (ω) = .85, indicating a good scale reliability.

2.3 Social Reinforcement Learning Task

To assess learning, we used a social reinforcement learning task with PLDs as feedback (see Figure 1). Participants were required to categorize randomly generated two-digit numbers (e.g. 99) into arbitrary groups ("A" or "B") via button press. Importantly, feedback was delivered probabilistically: in 85% of trials, correct responses were followed by rewarding feedback, while in 15% of trials correct responses were followed by non-rewarding feedback. This contingency was chosen to provide an appropriate level of difficulty for learning the underlying associations. Participants were informed that categories were arbitrary with no underlying rule, requiring them to learn via trial-and-error from feedback. The exact contingencies were not disclosed.

After each response, participants received PLD feedback. First, a random pattern of point lights was displayed, which transformed into either social or non-social feedback. In social blocks, point lights formed happy or angry human faces to indicate correct or incorrect responses. In non-social blocks, they formed check marks or crosses. Participants completed two blocks for each condition, each block consisting of six trials. Within each trial, 4 unique two-digit numbers were presented and repeated across subsequent trials, with the presentation order randomized. To avoid carry-over effects, even numbers were assigned to the social blocks, uneven numbers to the non-social blocks. Response options ("A" and "B") were displayed to the left and right of the screen, with their positions randomly switched between trials to avoid simple motor learning.

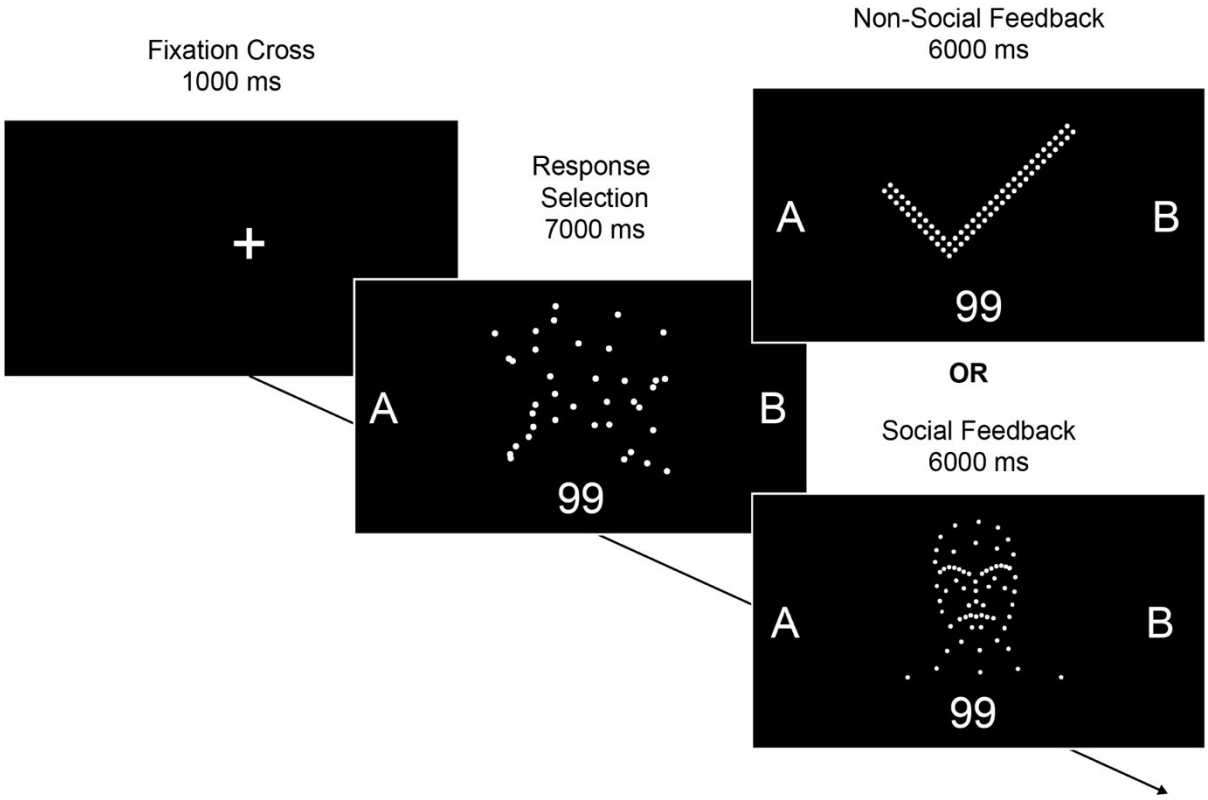

Figure 1: Illustration of a single trial. A random arrangement of point lights was displayed during response selection and subsequently transformed into either social or non-social feedback stimuli. Participants were asked to categorize a random number stimulus into one of two arbitrary groups, A or B. Non-social feedback consisted of videos of checkmarks and crosses (not displayed), while social feedback featured point-light faces expressing happy or angry emotions. Feedback was delivered probabilistically based on a reward contingency of 0.85/0.15. Each block consisted of six trials, with two social and two non-social blocks, for a total of 12 trials. Blocks were presented in an interleaved order, with the starting block type randomly selected. Note: Display brightness has been enhanced for visibility.

To control for order effects, the first block type (social / non-social) was randomly selected, and subsequent blocks alternated between the two types. The dependent variable for the GLMM analysis was the proportion of correct answers, defined as the proportion of trials, in which participants chose the high probability option. We expected participants to perform around chance level (50%) in the first trial and improve in subsequent trials, as they learned the underlying reward contingencies. A German version of the task is available online (Chrome recommended): https://raimund-buehler.github.io/SOCIALRL_PLD/

2.4 Statistical analysis

GLMM

We used a generalized linear mixed model (GLMM) to analyze learning performance. This approach accounts for repeated-measures and dependencies in the data, such as blocks and trials nested within individuals, by incorporating a random intercept for each individual (Lai & Kwok, 2015). The model includes both fixed effects, which capture overall trends in the sample (e.g., the average learning trajectory) and random effects, which capture between-person variability around these average relationships (e.g., individual differences in learning trajectories). This design allows us to examine individual learning trajectories and differences in performance across blocks. Specifically, we investigated the effect of AQ scores on participants' learning trajectories across block types, exploring whether autistic traits are associated with lower learning performance and whether this relationship occurs only in social blocks or across both block types.

Given the binary nature of the dependent variable, participant choice (correct/incorrect), a binomial distribution and logit link function was assumed. The following predictors were included in the model as fixed effects: AQ score (continuous), block type (binary: social / non-social) and trial number (continuous, 1 to 12). We performed grand-mean centering for AQ score by subtracting the overall sample mean from each individual's score (Bauer & Curran, 2005). This was done to improve model convergence and ensuring more stable parameter estimates.

To assess the presence of interindividual variability in trials and block types, we performed a likelihood ratio test, to evaluate whether the inclusion of random slopes significantly improved model fit. First, we compared a model with random slopes to a simpler model that included only random intercepts by subject and fixed effects, but no random slopes. The likelihood ratio test was significant ($\chi^2(5) = 160$, $p < 0.001$), confirming significant interindividual variance in learning trajectories and block types. To examine whether AQ scores accounted for this variability in learning trajectories, we added an interaction term, yielding the following model, with AQ specified as fixed effect, trial number and block type specified as both fixed and random effects. Specifically, in pseudo-code:

$$Choice \sim Block\ Type \times Trial\ Number \times AQ\ Score + (1 + Block\ Type + Trial\ Number \mid Participant) \quad (1)$$

All model estimates are reported as odds ratios (OR) incl. 95% confidence intervals (CIs) calculated using the Wald method. Odds ratios range from 0 to ∞, with values greater than 1 indicating increased odds of giving a correct answer and values less than 1 reflecting decreased odds. Confidence intervals that include 1 suggest no significant association, whereas intervals excluding 1 indicate a statistically significant effect. All models were fitted via the maximum likelihood estimation method using the *lme4* package (version 1.1-31) (Bates et al., 2015).

Reinforcement Learning Model

Probabilistic reinforcement learning tasks lend themselves to a trial-by-trial analysis of responses via reinforcement learning models, which are characterized by a set of parameters (Sutton & Barto, 2018). These parameters capture certain aspects of the individual learning process, i.e. the ability to quickly shift behavior in response to feedback (*learning rate*), the impact of reward reception (*reward sensitivity*), or the tendency to explore choice options under uncertainty (*temperature*). They in turn rely on the prediction error, a central component of reinforcement learning, which signals reward reception, but decays as reward contingencies grow to be expected.

We used a Rescorla-Wagner reinforcement learning model to analyze choice behavior in a trial-by-trial approach. Such models use a set of computational steps to predict choice probabilities given the current expected probability of receiving a reward and the choice outcome on the current trial. The first step consists of computing a reward prediction error (RPE) for a given trial *t*:

$$RPE(t) = \rho \times Outcome(t) - P_A(t) \quad (2)$$

With *P* denoting the estimated probability of receiving a reward and *A* denoting the choice option. The outcome is coded as a binary variable (0 for no reward and 1 for reward reception) and is scaled by the *reward sensitivity parameter* ρ. The RPE takes large values if a reward is received (positive) or omitted (negative) unexpectedly. Each trial, the expectation of receiving a reward on the next trial is updated by scaling the RPE by the *learning rate parameter* α:

$$P_A(t+1) = P_A(t) + \alpha \times RPE(t) \quad (3)$$

Here, the RPE is used to compute the expected reward probability when choosing option A on the next trial. If a reward was received unexpectedly, the RPE will take a large positive value and the expected probability of receiving a reward on the next trial will increase. Likewise, if a reward was omitted, the RPE will be negative and the expected probability of receiving a reward on the next trial will decrease. The result of this computation is scaled by the learning rate α, which is bounded

between 0 and 1 and can be estimated for each participant. Lastly, the probability of choosing the option is modeled using the softmax rule (a sigmoid function) and the *inverse temperature parameter* θ:

$$P_{Choice\ A}(t+1) = \frac{1}{1 + e^{-\theta(P_A - P_B)}} \quad (4)$$

The crucial point during this step is the scaling of the difference between expected probabilities for choices A or B. If θ is large, even a small difference in expected reward probabilities will bias choice behavior

towards A or B. Conversely, if it is small, the difference between probabilities needs to be very pronounced for choice behavior to be affected. This in turn results in more or less exploration of choice options, with more deterministic choice behavior for large values of θ and more probabilistic behavior for low values.

This set of equations was used to model individual choice behavior and estimate the described parameters using a set of custom python scripts. Parameters were estimated using maximum likelihood (ML), specifically by minimizing the negative likelihood using the *minimize* function from the *scipy* package with the TNC method (Virtanen et al., 2020). Parameters were initialized randomly with bounds manually constrained to -10 to 10 for α and ρ and 1 to 50 for θ. To ensure that parameters fell within the conventional bounds of 0 and 1, a logistic transformation was applied to α and ρ. For θ, no logistic transformation was applied. To further improve model fit, prevent extreme values and stabilize parameter estimation, we used elastic net regularization with L1 and L2 penalties with regularization strength set to $\lambda = 0.01$ for both penalties. This resulted in stable estimations over several cycles of parameter fitting. Parameters were estimated separately for wins and losses as well as social and non-social conditions, resulting in a set of 6 parameters for each participant.

As evident from histograms, scatterplots, as well as the fact that α and ρ are bounded between 0 and 1, assumptions for linear models may be violated (see below). Indeed, QQplots for linear models confirm deviation from normal distribution, mostly at the tails (see supplementary material, Figure S1). This suggests more extreme values in the observed data than expected under the assumption of normal distribution. Thus, for α and ρ, a GLMM with a beta distribution was considered an appropriate choice for analysis, as the beta distribution is defined on the interval between 0 and 1 and therefore naturally suitable to model data that is bounded within this range. θ, which was not bounded between 0 and 1, was investigated using a classical LMM approach. Diagnostic plots with the DHARMa package still indicated some assumption problems for the ρ_{win}-model in the beta-GLMM approach, however points aligned more closely to the diagonal line in the QQplot than for LMM models (see supplementary material, Figure S1 and S2).

All statistical analyses were performed using R 4.2.2 (R Core Team, 2022). LMM's and GLMM's were fit via maximum likelihood using the *lme4* package version 1.1.31 (Bates et al., 2015). Beta-GLMM's were fit using the *glmmTMB* package (Brooks et al., 2017) Assumptions for beta GLMM's were investigated using the *DHARMa* package version 0.4.6 (see supplementary material, Figure S2). Confidence intervals for odds ratios were computed using the Wald method.

In line with open science practices, Python and R scripts used for data pre-processing and statistical analysis are available online: https://github.com/raimund-buehler/RALT_final.nosync
The statistical analysis was pre-registered at the Open Science Framework (OSF): https://osf.io/tgcqm
A description of deviations from the preregistered analysis is provided in the supplementary material.

3 Results

Sample characteristics

63 individuals were selected for analyses in the GLM. 57.1% (N = 36) of the sample were females and 42.9% (n = 27) were male. The majority of the sample was composed of young adults (mean age = 26.51, SD = 6.76), with the oldest participant being 63 years old. With regards to AQ scores, the sample mean was M = 9.45, SD = 4.88 and values ranged from 1 to 25 (See Figure 2 **a**). Skewness of AQ scores was 0.82, indicating a moderately right-skewed distribution. This suggests that while most participants scored lower on the AQ, a smaller number of participants had higher scores, leading to a longer right tail.

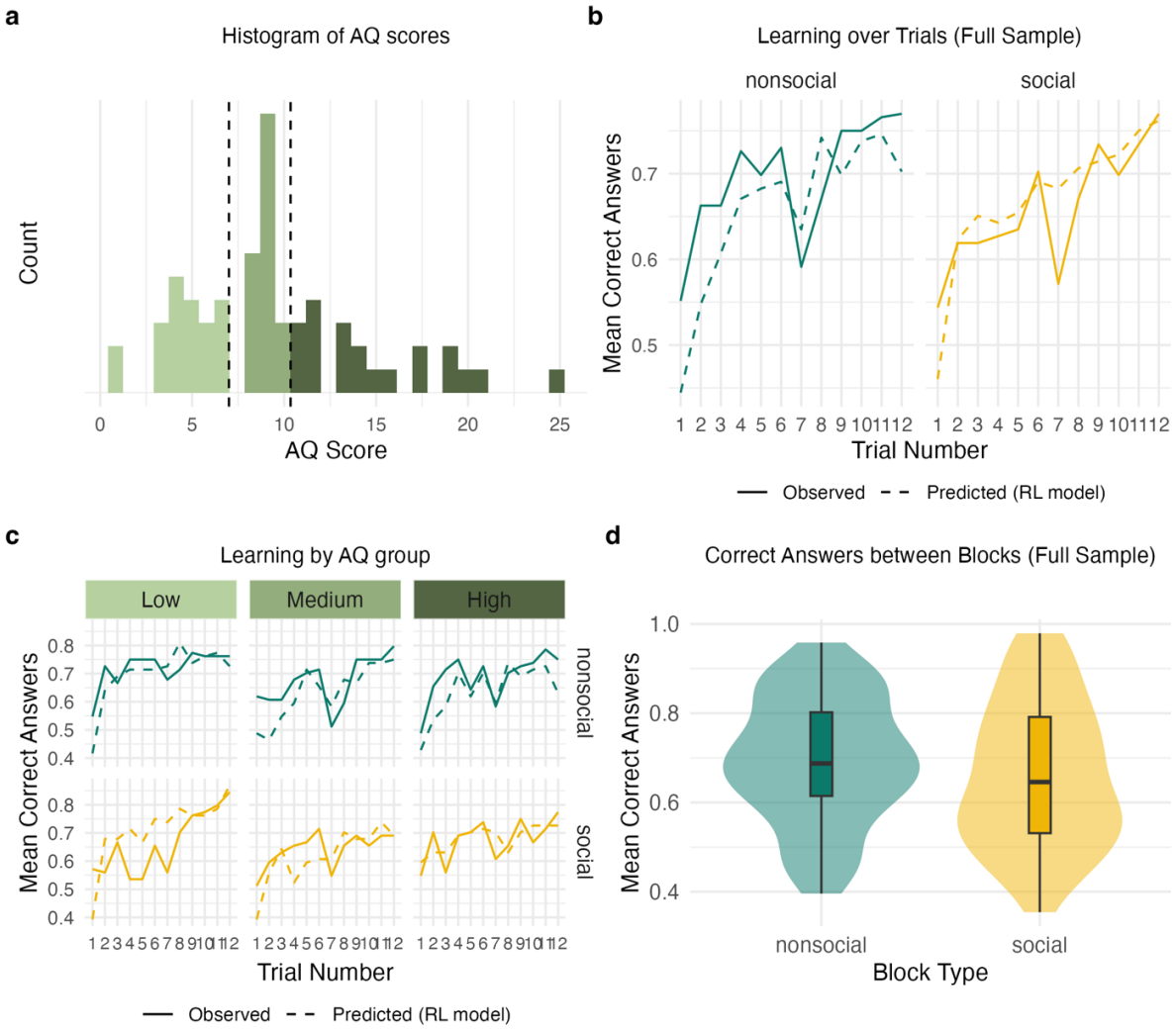

Figure 2 - a Histogram of AQ scores, split into 3 groups indicated by dashed lines (Low, Medium, High) **b** Learning over trials for all participants. Trial 1-6 were presented in the first block, trial 7-12 in the second block, resulting in a delay and a drop in performance in trial 7. Predicted learning performance of the reinforcement learning model is depicted as dashed lines and indicates satisfactory model fit **c** Learning over trials, split by 3 AQ groups depicted in the histogram (Low, Medium, High) for observed and predicted performance. Predicted performance of the RL model is depicted as dashed lines **d** Violin plot and boxplots for correct answers between blocks, averaged over all trials. Performance was marginally lower in social blocks across the full sample, but not significantly (p = 0.23).

GLMM

GLMM analyses revealed that trial number was significantly associated with higher odds of giving a correct answer OR = 1.10, 95% CI [1.06, 1.14], $p < 0.001$. This suggests that individuals learned the underlying association over trials, increasing their odds of giving a correct answer on average by 10% each trial. The main effect of block type (non-social vs. social), was not significant, OR = 0.81, 95% CI [0.61, 1.09], $p < 0.23$, although learning performance was marginally lower in social blocks (as suggested by an OR value below one).

The interaction between block type and AQ was not significant, OR = 1.03, 95% CI [0.97, 1.10], $p = 0.22$. This suggests that the effect of AQ did not differ significantly between block types. Similarly, the interaction between AQ score and trial number, the interaction between block type and trial number as well as the three-way interaction were not significant (all $p > 0.05$). This suggests that autistic traits did not explain variance in learning trajectories over trials, as indicated by largely parallel learning trajectories for different AQ groups (see Figure 3).

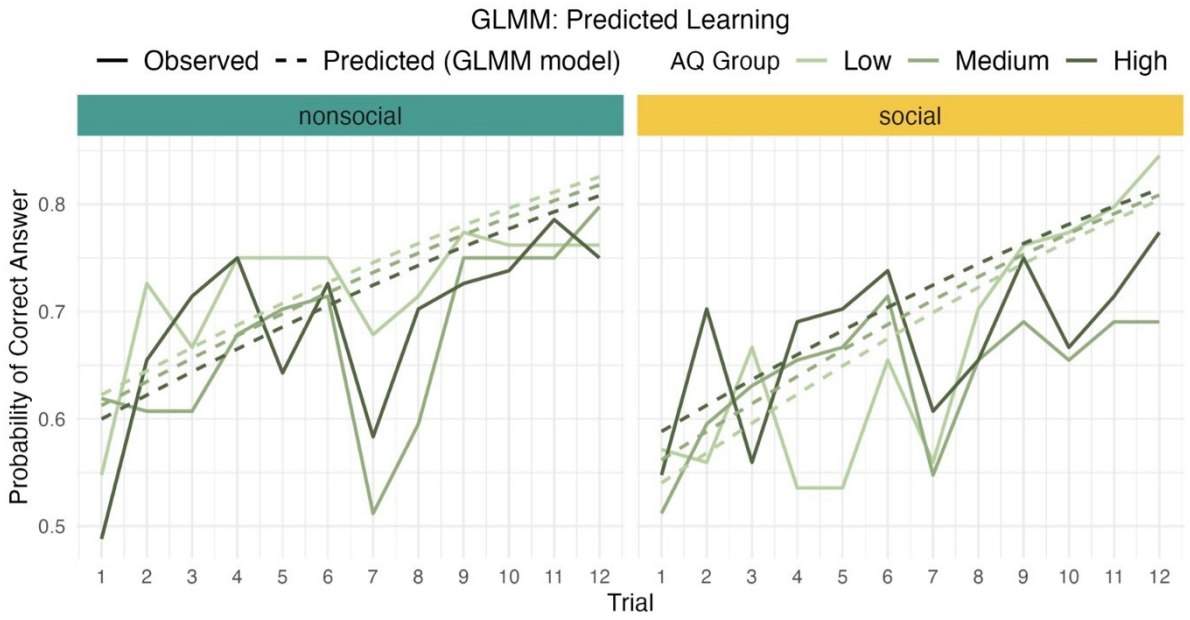

Figure 3: Observed (solid) and predicted (dashed) line plots, depicting learning over trials. Participants improve over trials, which is captured by the positive slope of model predictions. However, learning is not significantly different between block types or AQ groups (all $p > 0.05$).

Rescorla-Wagner Results

Model predictions and model fit for the estimated parameters can be investigated in Figure 2 **b** and **c** (dashed lines). Predictions follow the observed data closely, indicating satisfactory model fit. The histogram in Figure 4 **a** shows the distribution of fitted parameters over all participants. The non-parametric Wilcoxon signed-rank test suggested all parameters in the win domain to differ significantly from their counterparts in the loss domain (all $p < 0.001$), with parameters generally being larger for wins compared to losses.

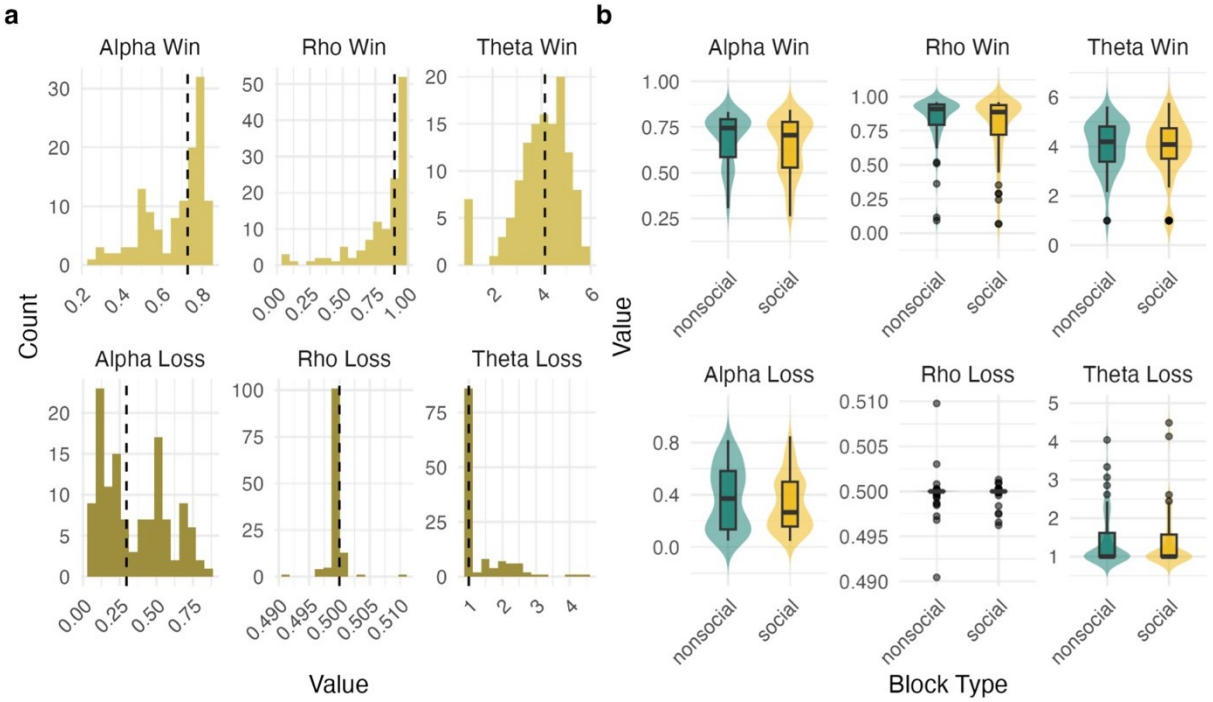

Figure 4: a Histograms of parameter distributions over all participants, separately for wins and losses. Dashed black lines depict median values. **b** Violin- and boxplots for all parameters, comparing social and non-social blocks.

When comparing parameters between social and non-social block types with random intercepts by subject, trends were observed for both α_{win} and ρ_{win} in a beta GLMM (see Figure 4 **b**). For ρ_{win}, the difference between social and non-social blocks was close to significant, OR = 0.75, 95% CI [0.56, 1.01], p = 0.0596, suggesting *reward sensitivity* for wins was marginally lower in social blocks. For α_{win}, the difference between block types was similar, OR = 0.83, 95% CI [0.68, 1.02], p = 0.0732, again suggesting that *learning rates* were marginally lower in social blocks. This aligns with the trend of reduced learning performance in social blocks described above, however these differences failed to reach significance. For θ, no difference between block types was observed.

To assess the effect of AQ, a separate model with AQ and block type as predictors was fit for each parameter and domain (win/loss). For the beta-GLMMs, the main effect of AQ was significant for the ρ_{win}-model, OR = 0.79, 95% CI [0.64, 0.98], p = 0.029 in a model specified without interaction. To further investigate the significance and robustness of this result, we performed two non-parametric tests, bootstrapping and a permutation test, both of which confirm the stability of the result, i.e. the empirical p-value derived from the permutation test was p = 0.003 (see supplementary material, Fig. S3).

In a model with an interaction term, AQ × block type was not significant (p = 0.74), suggesting no difference in the relationship between AQ and ρ_{win} across social and non-social blocks However, in a simple

slope analysis, the effect of AQ was significant in social blocks, OR = 0.95, 95% CI [0.90, 0.999], p = 0.046, but not in non-social blocks, OR = 0.96, 95% CI [0.91, 1.01], p = 0.11. Importantly, while this finding suggests that higher AQ scores were associated with a decrease in ρ_{win} in the social condition, it does not provide evidence that the relationship differs significantly between social and non-social blocks, as indicated by the non-significant interaction term. Thus, caution should be exercised when interpreting differences between conditions. No main effects of AQ or interactions of AQ × block type were observed for α_{win} or α_{loss} (see Figure 5, all p > 0.05).

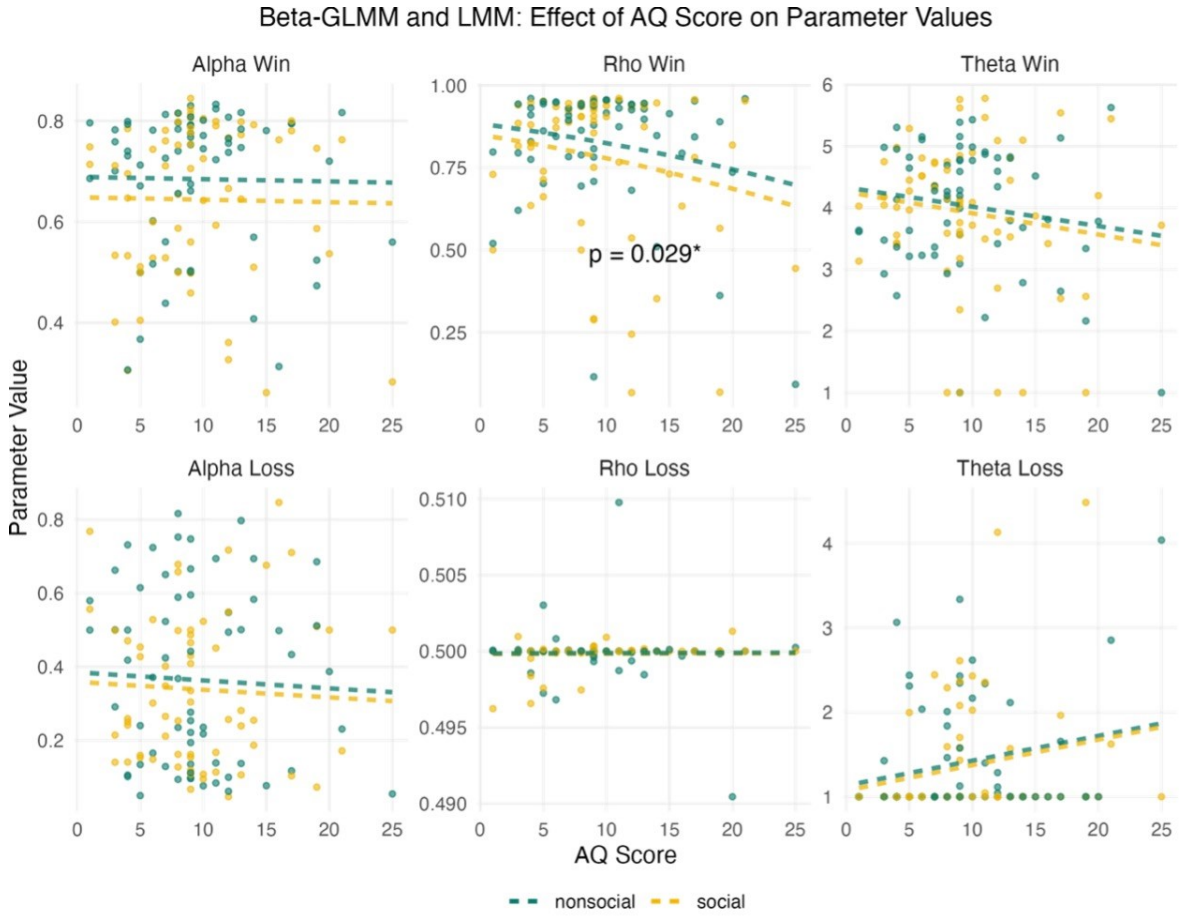

Figure 5 Scatterplots and line plots for the effect of autistic traits on parameter values in non-social and social blocks. Model predictions for each parameter are depicted as dashed lines. For alpha and rho values, a beta-GLMM was used, as parameters are bounded between 0 and 1. For theta values, an LMM was used, as parameter values are not bounded.

θ_{loss}, analyzed with a linear mixed model, was significantly positively associated with AQ, $\beta = 0.148$, $t = 2.23$, $p = 0.029$, $r = 0.27$. This relationship is also evident in Figure 5. However, as QQplots indicated that assumptions of linear models were violated for this model (see suppl. Fig. S1), results should be interpreted cautiously. The interaction of block type and AQ was not significant for the θ_{loss}-model (p > 0.5).

4 Discussion

To our knowledge, this study represents the first attempt to investigate the impact of autistic traits on learning from social and non-social PLDs matched on low-level visual properties. Our goal was to investigate raw learning performance, while also examining more subtle differences in the underlying learning processes by exploring parameters of a reinforcement learning model. These parameters were used to test competing theoretical accounts of ASD symptoms. In the analysis on raw learning performance, we did not detect significant differences related to autistic traits, however computational modeling did reveal attenuated reward sensitivity in response to wins for participants with high autistic traits. We will first discuss the analysis of raw learning performance and then move on to more intricate observations derived from the reinforcement learning model.

A critical result in the first analysis is that participants' performance improved over time in both social and non-social conditions, as evidenced by the significant effect of trial number. This illustrates that PLDs can successfully be used to investigate learning from social and non-social cues and emotional expressions depicted in facial PLDs have been recognized to an extent that supports learning. Regarding the main effect of condition (social vs non-social), learning performance was marginally lower in the social condition across the entire sample, but this difference was not significant. This would indicate that learning from social PLDs was slightly harder, possibly due to higher ambiguity in the social condition (i.e. non-social PLDs could easily be interpreted, whereas facial PLDs required recognizing depicted emotions from point light arrangements). However, this ambiguity did not significantly impact learning performance across the entire sample.

Regarding the effect of autistic traits, the main effect of AQ on learning performance was not significant, a result that is inconsistent with a general learning deficit for participants with high AQ, which would have manifested in decreased learning performance in both conditions. Contrary to our expectations, we also did not observe a significant interaction between condition and autistic traits, indicating that participants with high AQ scores did not perform worse in the social condition. As an explanation for the absence of this expected effect, we offer the following interpretation: participants with high AQ may have used *compensatory strategies* which did not rely on automatic emotion recognition, but rather on explicit cognitive effort. This compensatory strategy has been discussed in the literature as a potentially confounding factor in studies on facial emotion recognition and seems especially relevant in an adult sample of neurotypical individuals without ASD diagnosis (Harms et al., 2010). Furthermore, emotional expressions in facial PLDs were mostly discernible by focusing on the mouth, whereas less emotional information was conveyed by the eyes. Deficits in emotion recognition in ASD have been shown to relate to fixations to the eye region (Kliemann et al., 2012), thus this feature of the stimuli may have contributed to the missing effect. Furthermore, debriefing information suggested that participants experienced emotion recognition in social PLDs as particularly challenging. Reliance on abstract visual stimuli thus may have led to increased difficulty in the social condition for participants with low AQ and enhanced learning performance for those with high AQ. Altogether, this might have potentially masked interindividual learning differences.

Despite the absence of differences in raw learning performance, computational modeling was able to uncover more subtle differences in the learning process. In particular, the analysis using the Rescorla-Wagner model showed a trend for a main effect of condition on both *reward sensitivity* and *learning rate* parameters for the win domain (ρ_{win} and α_{win}). Both parameters were marginally lower in the social compared to the non-social condition over the entire sample, a finding that aligns with slightly lower learning performance in the social condition discussed above. Importantly, a significant main effect of AQ on ρ_{win} was observed, suggesting

attenuated sensitivity for rewarding feedback in participants with higher autistic traits. This suggests that the *internal value* assigned to rewarding feedback was lower in participants with high AQ.

We did not observe a main effect of AQ on learning rates, again suggesting no general learning deficit over conditions. We also did not observe an interaction of AQ and condition on learning rates, i.e. no differences in learning rates related to AQ between social and non-social conditions. Lastly, θ_{loss}, a parameter which captures the tendency to explore choice options in response to losses, was positively associated with AQ scores. This would suggest that participants with higher autistic traits exhibited more deterministic choice behavior and less exploration in response to losses. This finding can be interpreted to align with symptoms of repetitive behavior and insistence on sameness (Bishop et al., 2013), however assumptions for this analysis are likely not met (see supplementary QQplot, Fig. S1).

In summary, computational modeling indicates that participants with high autistic traits did not differ in terms of the weighting of prediction errors (neither social nor non-social), but rather in the way the *internal value* of stimuli was represented during learning. This aligns with the slow updating hypothesis (Lieder et al., 2019; Vishne et al., 2021), which suggests that rather than excessively depending on prediction errors, autistic individuals integrate recent information more gradually. Instead of rapidly adjusting beliefs based on new evidence, they rely more on longer-term accumulated knowledge. This is consistent with previous findings showing diminished trial-by-trial adjustments in both perceptual and motor learning tasks (Lieder et al., 2019; Vishne et al., 2021). In contrast to the volatility hypothesis (Lawson et al., 2017), which predicts excessive sensitivity to change, our results suggest that high-AQ individuals are less responsive to new feedback, supporting the idea that autistic learning difficulties might stem from underweighting recent experience.

Our results provide less support for the social motivation hypothesis (Chevallier et al., 2012), which suggests that autistic traits are linked to reduced learning specifically in social contexts. While we observed lower reward sensitivity in high-AQ participants, this effect was not stronger in the social condition. This suggests that autistic traits may affect learning more generally, rather than being driven by a diminished sensitivity to social rewards.

Some limitations of this study also need to be taken into account. First, observed differences in parameter values did not reflect in learning performance, i.e. we did not observe reduced learning for participants with high autistic traits. Thus, the effect of autistic traits on learning from PLDs was confined to the parameters of the reinforcement learning model. One could question whether a difference not reflected in learning performance is valid. However, we argue that differences in how associations are learnt, rather than exclusively focusing on learning outcome, is crucial for gaining a full picture of learning deficits in ASD. As mentioned above, compensatory mechanisms could have been employed by participants with high AQ to arrive at the same level of performance, however the process might differ.

Second, a potential confound in this study is the role of working memory (Collins & Frank, 2012). The task required participants to hold and manipulate information about probabilistic feedback, which places a significant demand on working memory resources. Differences in working memory capacity could influence participants' performance independently of their learning rates or reward sensitivities. We did not record measures of working memory in this study; however, we screened for intelligence deficits and did not include individuals with reduced cognitive abilities. Furthermore, as discussed above, a direct effect of AQ on learning performance was not observed. Thus, it seems unlikely that working memory would be a confound in these analyses. Still,

future studies should consider assessing and controlling for working memory capacity to disentangle its effects from those of the primary variables of interest.

Lastly, it is important to note that the study did not include participants with ASD but neurotypical controls from the healthy population. Although many findings on learning from social stimuli generated in the ASD population have been shown to generalize to autistic traits in neurotypicals (Bianco et al., 2020; Coll et al., 2020; Goris et al., 2021; Heerey, 2014; Karvelis et al., 2017; Sevgi et al., 2020), this limitation should be considered when extending our findings to individuals with ASD.

Conclusions

Despite limitations mentioned above, this study provides an important addition to the field of reinforcement learning in ASD, constituting the first attempt to employ carefully matched PLD stimuli in a reinforcement learning task. Results illustrate significant effects of autistic traits on different parameters of the learning process, providing a detailed description of learning deficits and their relation to theoretical models of ASD. Specifically, high autistic traits were related to reduced reward sensitivity in the win domain. In contrast, no effects of autistic traits were observed for learning rates, demonstrating the importance of distinguishing between different parameters of computational models. Future studies aiming to unravel distinctive contributions of learning rates and reward sensitivities could apply reinforcement learning from PLDs in conjunction with fMRI: PLDs are especially useful in situations where visual properties of the stimulus need to be fully accounted for. Activation in areas linked to face perception such as the FFA (Atkinson et al., 2012) or dissociable components of motivation, such as the ventral striatum and the ventromedial prefrontal cortex (Berridge et al., 2009) could then be related to perception of facial PLDs and model parameters. This could give further insight into social learning deficits in ASD and shed light on the root causes of observed social deficits.

Acknowledgments and Funding

We would like to thank the CoBeNe doctoral school and the University of Vienna for providing funds for this project, our master students who assisted in data collection (Nina Sireč and Serap Özlü), the CogSciHub-SPP team (David Cserjan) and all participants who took part in the study.

Conflict of Interest

The Authors declare no conflict of interest

References

American Psychiatric Association. (2013). DSM 5 Diagnostic and statistical manual of mental disorders. In Diagnostic and statistical manual of mental disorders: DSM-5 (Vol. 5). American psychiatric association.

Atkinson, A. P., Vuong, Q. C., & Smithson, H. E. (2012). Modulation of the face- and body-selective visual regions by the motion and emotion of point-light face and body stimuli. NeuroImage, 59(2), 1700–1712. https://doi.org/10.1016/j.neuroimage.2011.08.073

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Malesand Females, Scientists and Mathematicians. Journal of Autism and Developmental Disorders, 31(1), 5–17. https://doi.org/10.1023/A:1005653411471

Bassili, J. N. (1978). Facial motion in the perception of faces and of emotional expression. Journal of Experimental Psychology: Human Perception and Performance, 4(3), 373–379. https://doi.org/10.1037/0096-1523.4.3.373

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67, 1–48. https://doi.org/10.18637/jss.v067.i01

Berridge, K. C., Robinson, T. E., & Aldridge, J. W. (2009). Dissecting components of reward: "liking", "wanting", and learning. Current Opinion in Pharmacology, 9(1), 65–73. https://doi.org/10.1016/j.coph.2008.12.014

Bianco, V., Finisguerra, A., Betti, S., D'Argenio, G., & Urgesi, C. (2020). Autistic Traits Differently Account for Context-Based Predictions of Physical and Social Events. Brain Sciences, 10(7), Article 7. https://doi.org/10.3390/brainsci10070418

Bidet-Ildei, C., Decatoire, A., & Gil, S. (2020). Recognition of Emotions From Facial Point-Light Displays. Frontiers in Psychology, 11. https://doi.org/10.3389/fpsyg.2020.01062

Bishop, S. L., Hus, V., Duncan, A., Huerta, M., Gotham, K., Pickles, A., Kreiger, A., Buja, A., Lund, S., & Lord, C. (2013). Subcategories of Restricted and Repetitive Behaviors in Children with Autism Spectrum Disorders. Journal of Autism and Developmental Disorders, 43(6), 1287–1297. https://doi.org/10.1007/s10803-012-1671-0

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Machler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. The R Journal, 9(2), 378–400. https://doi.org/10.3929/ethz-b-000240890

Chambon, V., Farrer, C., Pacherie, E., Jacquet, P. O., Leboyer, M., & Zalla, T. (2017). Reduced sensitivity to social priors during action prediction in adults with autism spectrum disorders. Cognition, 160, 17–26. https://doi.org/10.1016/j.cognition.2016.12.005

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. Trends in Cognitive Sciences, 16(4), 231–239. https://doi.org/10.1016/j.tics.2012.02.007

Coll, M.-P., Whelan, E., Catmur, C., & Bird, G. (2020). Autistic traits are associated with atypical precision-weighted integration of top-down and bottom-up neural signals. Cognition, 199, 104236.

Collins, A. G. E., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. European Journal of Neuroscience, 35(7), 1024–1035. https://doi.org/10.1111/j.1460-9568.2011.07980.x

Dawson, G., Webb, S. J., & McPartland, J. (2005). Understanding the Nature of Face Processing Impairment in Autism: Insights From Behavioral and Electrophysiological Studies. Developmental Neuropsychology, 27(3), 403–424. https://doi.org/10.1207/s15326942dn2703_6

Ewbank, M. P., Pell, P. J., Powell, T. E., von dem Hagen, E. A. H., Baron-Cohen, S., & Calder, A. J. (2017). Repetition Suppression and Memory for Faces is Reduced in Adults with Autism Spectrum Conditions. Cerebral Cortex, 27(1), 92–103. https://doi.org/10.1093/cercor/bhw373

Freitag, C. M., Konrad, C., Häberlen, M., Kleser, C., von Gontard, A., Reith, W., Troje, N. F., & Krick, C. (2008). Perception of biological motion in autism spectrum disorders. Neuropsychologia, 46(5), 1480–1494. https://doi.org/10.1016/j.neuropsychologia.2007.12.025

Freitag, C. M., Retz-Junginger, P., Retz, W., Seitz, C., Palmason, H., Meyer, J., Rösler, M., & von Gontard, A. (2007). Evaluation der deutschen Version des Autismus-Spektrum-Quotienten (AQ)—Die Kurzversion AQ-k. Zeitschrift Für Klinische Psychologie Und Psychotherapie, 36(4), 280–289. https://doi.org/10.1026/1616-3443.36.4.280

Goris, J., Silvetti, M., Verguts, T., Wiersema, J. R., Brass, M., & Braem, S. (2021). Autistic traits are related to worse performance in a volatile reward learning task despite adaptive learning rates. Autism, 25(2), 440–451. https://doi.org/10.1177/1362361320962237

Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial Emotion Recognition in Autism Spectrum Disorders: A Review of Behavioral and Neuroimaging Studies. Neuropsychology Review, 20(3), 290–322. https://doi.org/10.1007/s11065-010-9138-6

Heerey, E. A. (2014). Learning from social rewards predicts individual differences in self-reported social ability. Journal of Experimental Psychology: General, 143(1), 332.

Hurlemann, R., Patin, A., Onur, O. A., Cohen, M. X., Baumgartner, T., Metzler, S., Dziobek, I., Gallinat, J., Wagner, M., Maier, W., & Kendrick, K. M. (2010). Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans. Journal of Neuroscience, 30(14), 4999–5007. https://doi.org/10.1523/JNEUROSCI.5538-09.2010

Huys, Q. J., Pizzagalli, D. A., Bogdan, R., & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: A behavioural meta-analysis. Biology of Mood & Anxiety Disorders, 3(1), 12. https://doi.org/10.1186/2045-5380-3-12

Jeste, S. S., & Geschwind, D. H. (2014). Disentangling the heterogeneity of autism spectrum disorder through genetic findings. Nature Reviews. Neurology, 10(2), 74–81. https://doi.org/10.1038/nrneurol.2013.278

Kaliukhovich, D. A., Manyakov, N. V., Bangerter, A., Ness, S., Skalkin, A., Boice, M., Goodwin, M. S., Dawson, G., Hendren, R., Leventhal, B., Shic, F., & Pandina, G. (2021). Visual Preference for Biological Motion in Children and Adults with Autism Spectrum Disorder: An Eye-Tracking Study. Journal of Autism and Developmental Disorders, 51(7), 2369–2380. https://doi.org/10.1007/s10803-020-04707-w

Karvelis, P., Seitz, A. R., Lawrie, S. M., & Seriès, P. (2017). Autistic traits, but not schizotypy, predict overweighting of sensory information in Bayesian visual integration (p. 230003). bioRxiv. https://doi.org/10.1101/230003

Keating, C. T., Fraser, D. S., Sowden, S., & Cook, J. L. (2022). Differences between autistic and non-autistic adults in the recognition of anger from facial motion remain after controlling for alexithymia. Journal of Autism and Developmental Disorders, 52(4), 1855–1871.

Kinard, J. L., Mosner, M. G., Greene, R. K., Addicott, M., Bizzell, J., Petty, C., Cernasov, P., Walsh, E., Eisenlohr-Moul, T., Carter, R. M., McLamb, M., Hopper, A., Sukhu, R., & Dichter, G. S. (2020). Neural Mechanisms of Social and Nonsocial Reward Prediction Errors in Adolescents with Autism Spectrum Disorder. Autism Research, 13(5), 715–728. https://doi.org/10.1002/aur.2273

Kliemann, D., Dziobek, I., Hatri, A., Baudewig, J., & Heekeren, H. R. (2012). The Role of the Amygdala in Atypical Gaze on Emotional Faces in Autism Spectrum Disorders. Journal of Neuroscience, 32(28), 9469–9476. https://doi.org/10.1523/JNEUROSCI.5294-11.2012

Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., & Jones, W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. Nature, 459(7244), 257–261. https://doi.org/10.1038/nature07868

Lai, M. H. C., & Kwok, O. (2015). Examining the Rule of Thumb of Not Using Multilevel Modeling: The "Design Effect Smaller Than Two" Rule. The Journal of Experimental Education, 83(3), 423–438. https://doi.org/10.1080/00220973.2014.907229

Lawson, R. P., Mathys, C., & Rees, G. (2017). Adults with autism overestimate the volatility of the sensory environment. Nature Neuroscience, 20(9), 1293–1299. https://doi.org/10.1038/nn.4615

Lenroot, R. K., & Yeung, P. K. (2013). Heterogeneity within Autism Spectrum Disorders: What have We Learned from Neuroimaging Studies? Frontiers in Human Neuroscience, 7, 733. https://doi.org/10.3389/fnhum.2013.00733

Lieder, I., Adam, V., Frenkel, O., Jaffe-Dax, S., Sahani, M., & Ahissar, M. (2019). Perceptual bias reveals slow-updating in autism and fast-forgetting in dyslexia. Nature Neuroscience, 22(2), 256–264. https://doi.org/10.1038/s41593-018-0308-9

Lin, A., Rangel, A., & Adolphs, R. (2012). Impaired Learning of Social Compared to Monetary Rewards in Autism. Frontiers in Neuroscience, 6. https://doi.org/10.3389/fnins.2012.00143

Mayer, J. L. (2017). The Relationship Between Autistic Traits and Atypical Sensory Functioning in Neurotypical and ASD Adults: A Spectrum Approach. Journal of Autism and Developmental Disorders, 47(2), 316–327. https://doi.org/10.1007/s10803-016-2948-5

Mottron, L., & Bzdok, D. (2020). Autism spectrum heterogeneity: Fact or artifact? Molecular Psychiatry, 25(12), Article 12. https://doi.org/10.1038/s41380-020-0748-y

Palmer, C. J., Lawson, R. P., & Hohwy, J. (2017). Bayesian approaches to autism: Towards volatility, action, and behavior. Psychological Bulletin, 143(5), 521–542. https://doi.org/10.1037/bul0000097

Palumbo, L., Burnett, H. G., & Jellema, T. (2015). Atypical emotional anticipation in high-functioning autism. Molecular Autism, 6(1), 47. https://doi.org/10.1186/s13229-015-0039-7

R Core Team. (2022). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing). https://www.R-project.org/

Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., & Baron-Cohen, S. (2015). Measuring autistic traits in the general population: A systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. Molecular Autism, 6, 2. https://doi.org/10.1186/2040-2392-6-2

Schuetze, M., Rohr, C. S., Dewey, D., McCrimmon, A., & Bray, S. (2017). Reinforcement Learning in Autism Spectrum Disorder. Frontiers in Psychology, 8. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.02035

Scott-Van Zeeland, A. A., Dapretto, M., Ghahremani, D. G., Poldrack, R. A., & Bookheimer, S. Y. (2010). Reward processing in autism. Autism Research: Official Journal of the International Society for Autism Research, 3(2), 53–67. https://doi.org/10.1002/aur.122

Sevgi, M., Diaconescu, A. O., Henco, L., Tittgemeyer, M., & Schilbach, L. (2020). Social Bayes: Using Bayesian Modeling to Study Autistic Trait–Related Differences in Social Cognition. Biological Psychiatry, 87(2), 185–193. https://doi.org/10.1016/j.biopsych.2019.09.032

Solomon, M., Frank, M. J., Ragland, J. D., Smith, A. C., Niendam, T. A., Lesh, T. A., Grayson, D. S., Beck, J. S., Matter, J. C., & Carter, C. S. (2015). Feedback-Driven Trial-by-Trial Learning in Autism Spectrum Disorders. American Journal of Psychiatry, 172(2), 173–181. https://doi.org/10.1176/appi.ajp.2014.14010036

Sowden, S., Schuster, B. A., Keating, C. T., Fraser, D. S., & Cook, J. L. (2021). The role of movement kinematics in facial emotion expression production and recognition. Emotion, 21(5), 1041–1061. https://doi.org/10.1037/emo0000835

Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: Predictive coding in autism. Psychological Review, 121(4), 649–675. https://doi.org/10.1037/a0037665

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., & Bright, J. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. Nature Methods, 17(3), 261–272.

Vishne, G., Jacoby, N., Malinovitch, T., Epstein, T., Frenkel, O., & Ahissar, M. (2021). Slow update of internal representations impedes synchronization in autism. Nature Communications, 12(1), 5439. https://doi.org/10.1038/s41467-021-25740-y

von der Lühe, T., Manera, V., Barisic, I., Becchio, C., Vogeley, K., & Schilbach, L. (2016). Interpersonal predictive coding, not action perception, is impaired in autism. Philosophical Transactions of the Royal Society B: Biological Sciences, 371(1693), 20150373. https://doi.org/10.1098/rstb.2015.0373

Zhang, L., Lengersdorff, L., Mikus, N., Gläscher, J., & Lamm, C. (2020). Using reinforcement learning models in social neuroscience: Frameworks, pitfalls and suggestions of best practices. Social Cognitive and Affective Neuroscience, 15(6), 695–707. https://doi.org/10.1093/scan/nsaa089

Supplementary Material

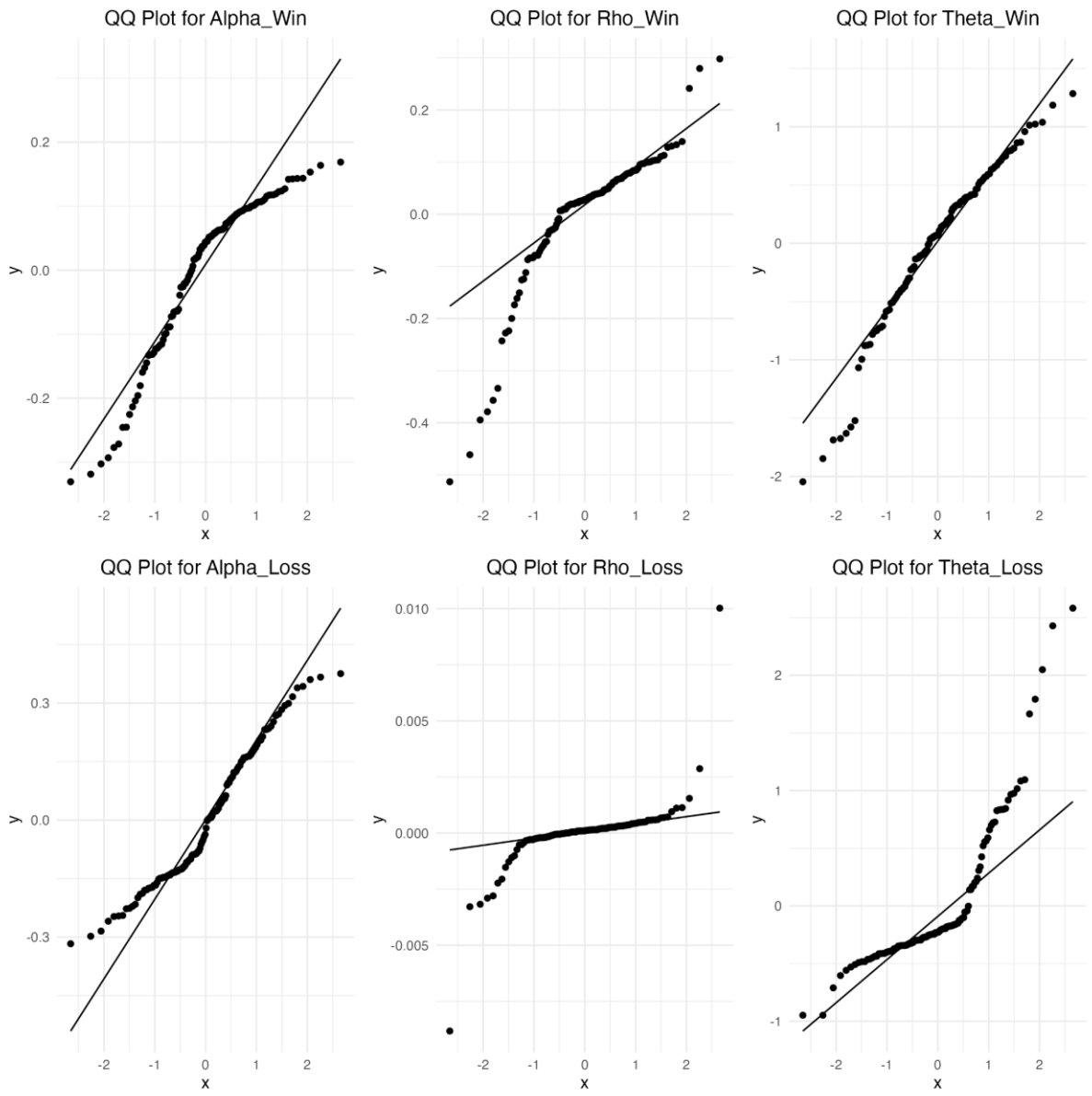

Figure S1 - QQplot for linear models fit on the RL-model parameters indicate deviation at the tails, suggesting more extreme values in the data than expected under the assumption of normal distribution.

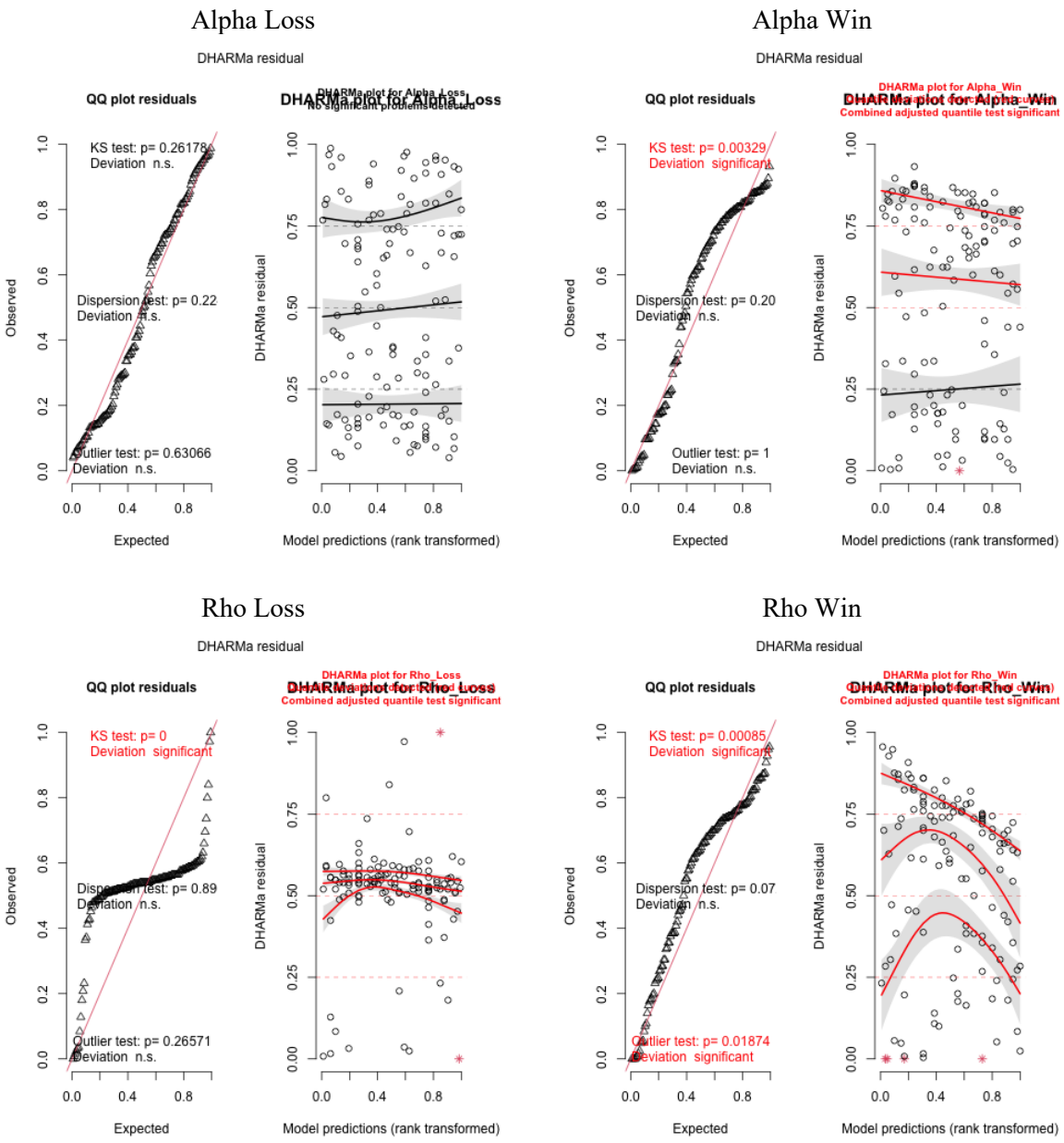

Figure S2 - Residual and diagnostic plots from the DHARMA package for beta-GLM models. Plots indicate some problems with deviation in the QQplots, however points align more closely than for the linear models. Additionally, the beta distribution is naturally defined on the interval between 0 and 1 and results from both LMM's and GLM's align regarding the main results.

Preregistration and Deviations from the Registered Analytical Plan

The preregistration was part of a larger project and thus included descriptions of an eye tracking task and analyses that are not directly relevant for this study. The initial intention was a joint analysis of eye tracking and reinforcement learning measures. The preregistration also mentions reaction time measures of the reinforcement learning task. After careful consideration of the complexity of analyses described in the main manuscript as well as the low additional theoretical and epistemic value of including further measures, the decision to omit these measures was made. Eye tracking and reinforcement learning data were split into two separate papers. Gender of subject was included as a predictor in exploratory models, but was not reported in the main manuscript, as it was not significant in any model ($p > 0.05$). The original term in the preregistration for "runs" is "cycles", however

both refer to the same concept, i.e. the consecutive presentation of the stimulus (see Fig. 1b in the main manuscript). The statistical analysis is described as a mixed ANOVA. Towards the end of the preregistration (under "exploratory analyses"), LMMs are mentioned as an alternative statistical approach to account for dependency in the data. However, as the dependent variable is bounded and can be interpreted as the number of failures and successes, a GLMM with a binomial distribution was considered a superior statistical approach and reported as such in the main manuscript. LMMs were also fit, however assumptions for these analyses were not well met (see QQplot, Fig. S1).

Bootstrapping and Permutation tests

To further investigate the significance and robustness of the effect of AQ on ρ_{win}, OR = 0.79, 95% CI [0.64, 0.98], p = 0.029, we performed two non-parametric tests, bootstrapping and a permutation test: Bootstrapping involves repeatedly sampling from the data (with replacement), fitting the model to each resampled dataset, and obtaining the parameter estimate (in our case, for AQ). This produces a normal distribution of estimates, allowing us to assess the robustness and variability of the parameter. If the original estimate lies near the center of the bootstrap distribution, it indicates small bias and robustness of the original estimate. Additionally, confidence intervals constructed from the bootstrap distribution provide a measure of uncertainty. If these intervals exclude zero, it suggests the estimate is statistically significant.

A permutation test, on the other hand, randomly shuffles the response variable, breaking any association between it and the predictors. By fitting the model to the shuffled data, we generate a distribution of estimates under the null hypothesis (centered around zero). The observed estimate is then compared to this null distribution, and an empirical p-value is calculated as the proportion of shuffled estimates that are as extreme as or more extreme than the observed estimate.

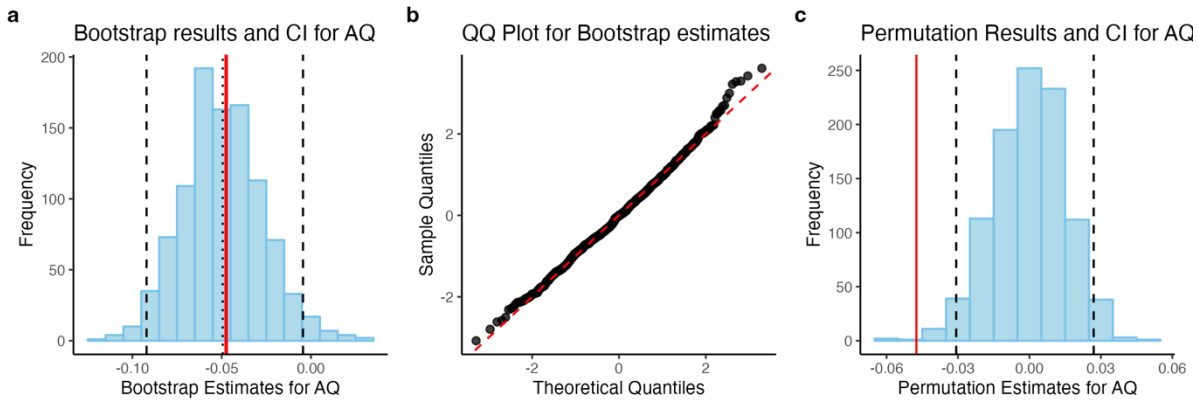

Figure S3 - a Bootstrap results. The observed estimate is depicted in red, mean and CI of bootstrapped estimates as dotted or dashed, black lines. The observed estimate aligns closely with the center of the bootstrap distribution, indicating small bias and robustness of the original estimate. CIs exclude zero **b** QQ-Plot for the bootstrap distribution. Points closely follow the diagonal line, suggesting bootstrap estimates follow a normal distribution. **c** Permutation results. Estimates for randomly shuffled data center around 0. The observed estimate (red line) lies outside the confidence intervals (dashed, black lines) for randomly shuffled data, indicating significance (empirical p = 0.003)

Chapter 3 – Interactive Effects of Oxytocin and Naltrexone on Social Gaze in Autism: Evidence from a Naturalistic Eye Tracking Study

Raimund Buehler[1], Ulrich Ansorge[2,3,4], Matthäus Willeit[5], Giorgia Silani[1,3]

[1]Department of Clinical and Health Psychology, Faculty of Psychology, University of Vienna, Austria

[2]Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, Vienna, Austria

[3]Vienna Cognitive Science Hub, University of Vienna, Vienna, Austria

[4]Research Platform Mediatized Lifeworlds, University of Vienna, Vienna, Austria

[5]Department of Psychiatry and Psychotherapy, General Hospital Vienna, Vienna, Austria

Corresponding authors:

Raimund Buehler
raimund.buehler@univie.ac.at
+43-1-4277-47241
Liebiggasse 5, 1010 Vienna, Austria

Giorgia Silani
giorgia.silani@univie.ac.at
+43-1-4277-47223
Liebiggasse 5, 1010 Vienna, Austria

Abstract

Social gaze, particularly eye contact, is essential for navigating human interactions and is commonly reduced in autism spectrum disorder (ASD). However, the neurobiological mechanisms underlying this reduction are not fully understood. Both oxytocin and opioids have been implicated in modulating social attention, potentially through interactive effects. While prior research has primarily focused their effect on social reward processing, stress-related social avoidance, implicating the hypothalamic-pituitary-adrenal (HPA) axis, may also play a critical role. In this study, we investigated social gaze during a naturalistic conversation in N = 73 participants (34 ASD, 39 controls) across four pharmacological conditions: naltrexone, oxytocin, their combination and placebo. Eye tracking data was collected during a face-to-face interview and salivary cortisol samples served as a marker of HPA axis activity. As expected, individuals with ASD fixated less on the eye region (d = -0.24) and more on the background (d = 0.27). Medication effects were more complex: while no drug x group interaction effects were observed, exploratory analyses revealed that between-group differences in eye-directed gaze were attenuated under pharmacological modulation. Importantly, naltrexone elicited a robust cortisol response that was differentially related to gaze patterns across groups: In the ASD group, higher cortisol was associated with *more* fixations on the eyes (r = 0.17), whereas in controls, higher cortisol related to *fewer* fixations on the eyes (r = -0.21). These findings suggest different effects of arousal or stress-related mechanisms under opioid antagonism across groups, possibly explaining the attenuated group differences. Our findings provide novel insight into how opioids and oxytocin interact with stress-related mechanisms to shape social gaze in ASD during naturalistic interactions.

1 Introduction

Social attention is a central aspect of human interaction. In particular, eye contact is among the most essential social signals for initiating and sustaining dyadic conversations. A broad range of studies has demonstrated that autism spectrum disorder (ASD) is associated with atypical gaze patterns, particularly reduced fixations on the eye region (Chevallier et al., 2015; Chita-Tegmark, 2016). In an effort to explain these differences, several candidates have been discussed as potential neurobiological substrates of atypical social attention in ASD.

Early accounts have emphasized the role of endogenous opioids in reinforcing social and bonding behavior, with social interactions assumed to stimulate opioid release (Panksepp, 1979). This perspective, known as the brain-opioid theory of social attachment (BOTSA), originated from findings in rodents, where social isolation produced distress behaviors reminiscent of opioid withdrawal that were mitigated by morphine administration (Kalin et al., 1988). In contrast, opioid antagonists such as naltrexone increased social seeking behavior (Fabre-Nys et al., 1982; Keverne et al., 1989).

Since the initial conceptualization of BOTSA, decades of research in rodents and primates have demonstrated the involvement of opioids in affiliative social behavior and attachment (Machin & Dunbar, 2011; Pellissier et al., 2018). In humans, opioids have similarly been established as a critical modulator of social behavior across multiple levels of analysis. For example, positron emission tomography (PET) studies have shown that social touch increases opioid-receptor availability in regions such as the striatum, insula, as well as frontal and cingulate cortices (Nummenmaa et al., 2016). Genetic evidence has linked the A118G variant of the OPRM1 gene, which encodes the μ-opioid receptor, to a greater sensitivity to social reward, affectionate relationships, and social rejection (Troisi et al., 2011; Way et al., 2009). Pharmacological studies have further implicated opioids across a diverse range of measures, including reports of social connection and attention to the eye region (Chelnokova et al., 2016; Løseth et al., 2024; Pellissier et al., 2018).

Results demonstrating the involvement of opioids in social behavior have prompted a long line of research on the effects of opioid antagonists in ASD. The original hypothesis derived from BOTSA proposed a permanently elevated opioid tone in individuals with ASD, which sparked a series of studies investigating whether naltrexone would normalize opioid activity and restore adaptive social behavior in ASD. Despite some evidence indicating benefit from low-dose naltrexone in ASD with elevated β-endorphin levels (Bouvard et al., 1995; Leboyer et al., 1992), clinical trials have generally reported limited efficacy of naltrexone in alleviating social symptoms (for a review see Roy et al., 2015) and current evidence to support its use for ASD treatment remains insufficient.

To date, the precise mechanism by which opioids regulate social behavior are still elusive and evidence regarding the effects of opioid antagonists such as naltrexone is conflicting (Putnam & Chang, 2022c). Specifically, it is not fully understood whether naltrexone increases or reduces affiliative social behaviors. According to BOTSA, opioid antagonism increases social seeking, a notion that has been supported by results of increased grooming in primates (Fabre-Nys et al., 1982; Keverne et al., 1989). In line with this, naltrexone heightens attention to emotional facial expressions in some studies (Wardle et al., 2016). However, *reduced* grooming in primates (Martel et al., 1993), as well as reduced social attention and social connection in humans has been reported as well (Chelnokova et al., 2016; Inagaki et al., 2016).

Several attempts have been made to reconcile such inconsistencies in theories emphasizing the *context* of opioid manipulation. For example, Løseth et al. propose a state-dependent model (SOMSOM), which posits that the behavioral effects of opioid antagonists depend on the individual's initial motivational state, enhancing social

motivation during distress by amplifying social seeking, but reducing social exploration in non-distressed states by dampening social reward sensitivity. Similarly, Pellissier et al. propose a μ-receptor balance model in which optimal social behavior relies on an equilibrium between μ-opioid activity and the social avoidance system (SAS), which (among other processes) encompasses the HPA axis. Consequently, both opioid hypoactivity (favoring SAS dominance and social withdrawal) and hyperactivity (leading to reward saturation and social indifference) may result in maladaptive social behavior.

More recently, increased attention has been devoted to other neuromodulators considered potential targets for ASD treatment, particularly oxytocin (Young & Barrett, 2015). Oxytocin is a nonapeptide which is synthesized in the hypothalamus and released into systemic circulation from the pituitary, acting both as a hormone and a neuromodulator. Effects of oxytocin on attachment and social behavior were first implied by results on pair bonding in prairie voles and subsequent findings of increased trust and social attention in neurotypical individuals (Guastella et al., 2008; Kosfeld et al., 2005; Young & Wang, 2004). Evidence implicating oxytocin in ASD emerged from studies showing atypical oxytocin levels in plasma of autistic children (Modahl et al., 1998) along with research linking variants of the oxytocin receptor gene to social-communicative impairments (S. G. Gregory et al., 2009; LoParo & Waldman, 2015). Increased social gaze in ASD following oxytocin administration (Auyeung et al., 2015) has been associated with increased social salience or decreased social stress (Shamay-Tsoory & Abu-Akel, 2016). However, it is important to note that studies with larger sample sizes have sometimes failed to replicate earlier effects (Guastella et al., 2015; Ooi et al., 2016) and the current literature on the social effects of oxytocin remains mixed (Quintana et al., 2021).

In summary, both opioids and oxytocin have been proposed as potential neurochemical mechanisms regulating social affiliation, but their exogenous manipulation has led to inconsistent findings both in neurotypical and ASD individuals. To resolve such inconsistencies, some researchers have recently suggested moving beyond the isolated investigation of neurochemical systems towards a more exhaustive account of their interactions and potential underlying mechanisms (Putnam & Chang, 2022c). For example, the interaction of oxytocin and opioids has been recognized for a long time (Vuong et al., 2010). A series of experiments in rodents provided evidence that opioid activity inhibits oxytocin and demonstrated that naloxone increased plasma oxytocin levels and firing-rates of oxytocin neurons (Bicknell & Leng, 1982; Vandesande & Dierickx, 1975). These findings suggest that the disinhibition of oxytocin by opioid antagonists may exert interactive effects on social behavior. Notably, such effects have been recently observed in non-human primates, where a combination of oxytocin and naloxone increased fixations on the eye region of conspecifics (Dal Monte et al., 2016), suggesting that a combined administration may help to alleviate some of the social difficulties observed in ASD (Fan et al., 2020). However, empirical evidence on this assumption is currently lacking.

In order to translate these findings to humans, we therefore conducted the first within-subject, double-blind, placebo-controlled pharmacological study in neurotypical and ASD participants to assess the effect of a combined administration of oxytocin and naltrexone on social attention during a naturalistic, face-to-face interaction.

Additionally, to elucidate potential underlying mechanisms of action, we explored the role of the HPA axis under combined pharmacological challenge, as it is known that both opioids and oxytocin interact with it, which in turn influences social behavior (Sandi & Haller, 2015). For example, elevated cortisol following social stress has been related to reduced social gaze in neurotypical males (Vatheuer et al., 2021) while reduced social

gaze in ASD has been attributed to both hypo- and hyperarousal (Cuve et al., 2018), possibly because it is experienced as aversive or uncomfortable, or due to a reduced sensitivity to the social cues (Moriuchi et al., 2017). In line with this, both hypo- as well as hyper-responsiveness of the HPA axis in ASD have been reported (Taylor & Corbett, 2014).

Hypotheses and Preregistration

Analyses and hypotheses were preregistered before data collection (OSF: https://osf.io/c3mkg). We hypothesized that participants with ASD would show reduced attention to the eye region compared to the neurotypical control (CTRL) group. Building on previous research conducted in primates, rodents and humans, we further expected that oxytocin (OXT) would increase fixations on the eye region compared to placebo across groups. As discussed above, previous research has provided conflicting evidence regarding the effect of naltrexone on social attention (NAL). Therefore, while we expected NAL to increase social attention across groups, alternative outcomes, including potential interactions with group, were considered plausible. Lastly, the combined administration of OXT and NAL was expected to exert a stronger effect on social attention than either drug administered in isolation.

In addition to these preregistered hypotheses, analyses on cortisol levels and their effect on social attention were added as exploratory analyses. We expected oxytocin to reduce salivary cortisol levels in both groups due to its inhibitory effect on the HPA axis (Heinrichs et al., 2003). In contrast, naltrexone was expected to elicit a cortisol response via disinhibition of the HPA axis (King, 2002). We made no predictions regarding the combination of both drugs, as reports of a combined administration are rare. Similarly, we made no predictions regarding group differences in cortisol levels, as both hypo- and hyperactivation of the HPA axis in ASD have been described (Taylor & Corbett, 2014). Lastly, we explored the relationship between cortisol reactivity and social attention. Specifically, we investigated how changes in cortisol levels were associated with fixations on the eye region. Based on the notion of hyperarousal-induced gaze avoidance (Stuart et al., 2022), we hypothesized that higher cortisol reactivity would be associated with reduced attention to the eye region in ASD. Alternatively, in line with the hypoarousal hypothesis (Moriuchi et al., 2017), it was also considered plausible that increased social gaze might be associated with heightened cortisol reactivity.

2 Methods

Sample

Participants for the neurotypical control group (CTRL) were recruited via the University of Vienna's participant management platform, while individuals with an ASD diagnosis were invited via outreach to specialized institutions and social media postings. Participants in the ASD group were required to submit a diagnosis by a licensed psychiatrist or psychotherapist. A total of 247 individuals completed a screening survey. For the control group, inclusion criteria required participants to be over 18 years old, fluent in German, with no vision impairment, no history of psychiatric or neurological disorders (or medication for such conditions), no traumatic head injury, no drug or alcohol addiction and no contraindication for oxytocin or naltrexone. The same criteria applied to ASD participants, although they were not excluded for comorbid disorders and medication use, as both are very common in the ASD population. The type of disorder and medication intake was recorded. Female participants were asked to provide information on contraceptive use, menstrual cycle history, and cycle length to be used as control variables.

After prescreening, 79 participants were recruited (39 ASD, 40 CTRL). Of these, 59 participants completed all four testing sessions (27 ASD, 32 CTRL). In the ASD group, 23 participants reported comorbid diagnoses, with depression (n = 12) and attention deficit hyperactivity disorder (ADHD; n = 10) most frequently reported. Correspondingly, the most common medications included selective serotonin/noradrenaline reuptake inhibitors (SSRIs/SNRIs; n = 19), stimulants (n = 8), and neuroleptics (n = 6). In total, 20 ASD participants reported regular medication use, with the number of medications ranging from one to four per individual.

18 participants dropped out prematurely (10 ASD, 8 CTRL), corresponding to a dropout rate of 23%. Of these, 16 individuals (9 ASD, 7 CTRL) discontinued after the first session, while two participants (one ASD, one CTRL) completed two sessions before dropping out. The most common reason for dropout were missed or canceled appointments (n = 14). These individuals' data were retained for analysis if possible. Four participants (3 ASD, 1 CTRL) reported side effects and chose to discontinue, all of whom were female and had received naltrexone. Two male participants were excluded from analysis, one due to low cognitive functioning (IQ = 79, ASD group) and another due to a positive opiate drug test (CTRL group). Data from participants who reported side effects were not included in the analyses.

A total of 254 testing sessions were conducted, including dropouts. For final analyses, data from 73 participants across 248 sessions were included, with six sessions excluded due to the above-mentioned reasons. Further demographic and questionnaire data from the 73 participants included in the analysis are summarized in Table 1.

A chi-square test for gender distribution showed no significant difference between groups in the final sample (p = 0.102, see Table 1). Of those participants who completed all sessions, marginally more females were in the ASD group (15 F, 12 M) and more males in the CTRL group (18 M, 14 F), however, this difference was again not significant ($\chi^2(1) = 0.41$, $p = 0.52$). One participant in the ASD group identified as non-binary and was excluded when the analysis involved gender. Ages ranged from 18 to 70 ($M = 31.36$, $SD = 12.10$) with a marginal age difference between groups, which was not statistically significant ($t(62) = 1.78$, $p = 0.080$, see Table 1). Verbal IQ, assessed with the Mehrwachwahl-Wortschatz-Intelligenztest (Lehrl, 1999), was not significantly different between groups ($t(54) = 1.45$, $p = 0.154$). Consistent with ASD diagnoses, significant differences were observed

for autism spectrum quotient (AQ) and the social interaction anxiety scale (SIAS; (Heimberg et al., 1992). Scores on the interpersonal reactivity index (IRI; (Davis, 1980) and its subscales were significantly different between groups except for the "Empathetic Concern" and the "Personal Distress" subscales (see Table 1). Scores on the state-trait anxiety inventory (STAI-t; (Spielberger et al., 1971) were significantly different ($t(56) = 2.84$, $p = 0.006$), indicating higher anxiety levels in the ASD group, consistent with individual reports of comorbid generalized anxiety disorder in the ASD group (2 mentions). In contrast, despite individual reports of comorbid depression, the group difference for the Beck depression inventory (BDI; Beck et al., 1978) was not significant, albeit approaching significance ($t(55) = 1.81$, $p = 0.075$). Generally, more negative and less positive affect was observed in the Positive and Negative Affect Schedule (PANAS; (Crawford & Henry, 2004) in the ASD group in earlier sessions. This difference appears to have subsided in later sessions. Sympathy ratings for the experimenter showed no significant differences across any dimension.

Experimental Procedure

The study was designed as a within-subject, double-blind, placebo-controlled trial with four sessions per participant. Each session lasted approximately 2.5 to 3 hours. Upon arrival, participants provided written informed consent and were introduced to the study protocol. To assess baseline mood, participants completed the Positive and Negative Affect Schedule (PANAS) at the start of each session. Immediately after, a urine test was conducted to screen for opioid use and pregnancy. If results were negative, the session continued with the collection of a baseline saliva sample. In the first session, verbal IQ was assessed using the Mehrwachwahl-Wortschatz-Intelligenztest (MWT).

Each session involved the administration of both a pill and a nebulizer solution, which contained either active medication or a placebo. To account for differences in drug onset, the pill was administered 50 minutes, the nebulizer 20 minutes prior to testing (see pharmacological procedure for more details). The nebulizer was applied for 5 minutes on each nostril (10 minutes total).

Variable			ASD		CTRL			
			N		N		$\chi 2$	p
Gender	Male		13		24			
	Female		20		15		2.68	0.102
	Non-Binary		1		0			
			Mean	**SD**	**Mean**	**SD**	**t**	**p**
Age			34.06	13.37	29.00	10.48	1.78	0.080
AQ - Autism Spectrum Quotient			24.31	5.90	9.18	4.41	11.36	**< 0.001**
AQ - Communication and Reciprocity			7.55	1.80	3.47	1.80	8.97	**< 0.001**
AQ - Social Interaction and Spontaneity			8.55	2.57	3.06	2.50	8.56	**< 0.001**
AQ - Imagination			8.21	2.62	2.65	1.67	9.84	**< 0.001**
AQ - Social Scales			16.10	3.88	6.53	3.65	10.04	**< 0.001**
BDI-A - Beck Depression Inventory			6.93	2.94	5.70	2.29	1.81	0.075
Verbal IQ (MWT)			112.59	17.15	107.67	10.70	1.45	0.154
IRI - Interpersonal Reactivity Index			61.26	12.57	70.71	10.80	-3.24	**0.002**
IRI - Empathetic Concern			17.45	5.30	19.50	5.09	-1.59	0.118
IRI - Fantasy			14.32	5.42	20.12	3.88	-4.91	**< 0.001**
IRI - Personal Distress			14.42	4.99	12.21	4.91	1.80	0.077
IRI - Perspective Taking			15.06	5.40	18.88	4.93	-2.97	**0.004**
SIAS - Social Interaction Anxiety Scale			51.90	11.07	25.59	14.03	8.37	**< 0.001**
STAI-t - State-Trait Anxiety Inventory (trait)			49.10	5.57	45.49	4.55	2.84	**0.006**
PANAS - Positive and Negative Affect Schedule								
	Negative Affect	**Session 1**	1.62	0.44	1.40	0.33	2.39	**0.020**
		Session 2	1.57	0.43	1.38	0.30	1.93	0.059
		Session 3	1.42	0.44	1.43	0.51	-0.08	0.933
		Session 4	1.54	0.50	1.38	0.42	1.31	0.196
	Positive Affect	**Session 1**	2.78	0.59	3.26	0.56	-3.55	**< 0.001**
		Session 2	2.64	0.68	3.21	0.69	-3.23	**0.002**
		Session 3	2.62	0.86	3.04	0.77	-1.93	0.059
		Session 4	2.75	0.85	2.88	0.75	-0.63	0.528
Sympathy Ratings	Interesting		4.78	1.46	5.14	1.47	0.88	0.381
	Good		5.84	1.12	6.21	0.82	1.98	0.051
	Kind		6.22	0.94	6.46	0.75	1.53	0.129
	Open-Minded		5.86	1.22	5.97	1.13	0.19	0.847

Table 1 - Demographic and questionnaire data. T- and p-values were obtained through standard t-tests, except for sympathy ratings, which were obtained using a linear mixed model with random intercepts by subject, as data was collected each session (within-subject). Questionnaires and subscales marked in bold highlight significant differences between groups.

Drug administration was followed by a waiting period, during which the experimenter set up the eye tracking equipment. Shortly before testing, at the estimated peak plasma levels for both drugs, participants completed a brief side effects scale, followed by a second saliva sample. Participants then performed a semi-structured interview, during which eye gaze was recorded using a mobile eye-tracking headset.

After task completion, participants were asked to guess which medication they had received to assess the success of blinding. They then provided sympathy ratings for the experimenter, using a 1-7 Likert scale. Lastly, a third and final saliva sample was collected. In the first session, participants also completed a battery of self-report questionnaires, including the Interpersonal Reactivity Index (IRI), the Social Interaction Anxiety Scale (SIAS), the State-Trait Anxiety Inventory (STAI-t), and the Autism Spectrum Quotient (AQ). Each session concluded with a short debriefing, during which participants received further information and contact details.

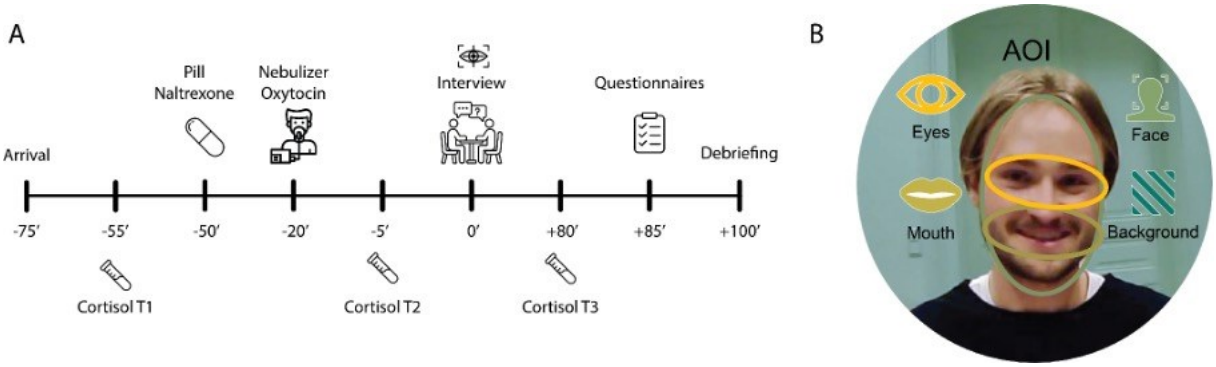

Figure 1 - Overview of the experimental design and eye-tracking analysis. A Timeline of the experiment: Participants always received a pill and a nebulizer solution, resulting in four pharmacological conditions: Naltrexone + Oxytocin (NAL/OXT), Naltrexone + Placebo (NAL/PLA), Placebo + Oxytocin (PLA/OXT), or Placebo + Placebo (PLA/PLA). Following drug administration, participants completed a semi-structured interview while wearing a mobile eye-tracking headset. Cortisol samples were taken at 3 timepoints. **B** Eye-tracking data was analyzed based on predefined Areas of Interest (AOIs), including eyes, mouth, face, and background.

Eye-Tracking Interview

A team of seven trained experimenters conducted the interviews, ensuring that each participant was consistently paired with the same experimenter across all sessions. To reduce potential differences in gaze behavior based on the sex of the interaction partner, participants were always interviewed by an experimenter of the same sex. The interviews took place in a quiet lab setting, where both the participant and the experimenter were seated approximately one meter apart at a table. Each session followed a semi-structured interview format, during which 13 scripted questions were read aloud by the experimenter from a printed list. These questions were designed to simulate natural conversation and were adapted from previous studies investigating social attention during face-to-face interactions (Auyeung et al., 2015). The questions were categorized into two types: closed questions (1–6), simple yes/no questions requiring minimal elaboration, and open questions (7–13), designed to encourage longer responses and greater social engagement. To maintain a natural flow of conversation, experimenters were permitted to use brief follow-up remarks, but were instructed to ensure the total interview duration remained between 7 and 10 minutes. Throughout the interview, the duration of each question was recorded via button press to allow for a time-course analysis of gaze behavior. A full list of interview questions is provided in the supplementary material.

The average interview duration was M = 9.47 minutes, SD = 2.82 minutes. A difference in interview duration between groups was observed, with slightly shorter interviews in the ASD group (*M* = 9.00 minutes, *SD*

= 2.50 minutes) compared to the CTRL group (M = 9.89 minutes, SD = 3.02 minutes; $t(247) = -2.52$, $p = 0.012$). However, since the dependent variable in the analysis (fixations per second) was normalized for the total interview duration, variations in interview length are unlikely to influence results.

Eye-Tracking Equipment and Procedure

The eye tracking procedure follows our previous work and is described in more detail there (Buehler et al., 2024). Eye-tracking data was collected using the Pupil Labs Core system (Pupil Labs, Berlin, Germany), a mobile head-mounted eye tracker with two infrared pupil cameras and a forward-facing world camera. Pupil cameras captured eye movements at a 120 Hz sampling rate. To estimate gaze direction, we used an approach based on a 3D model of the eye (Kassner et al., 2014), which accounts for head movement and slippage of the headset, improving the stability of gaze estimations. Participants were instructed to minimize large head movements, though natural conversational behaviors such as nodding or shaking the head were permitted.

Before starting the interview, a single-marker calibration choreography was performed. A printed calibration marker was held at eye level in front of the experimenter's face. The manufacturer reports a gaze estimation accuracy of approximately 0.6° visual angle (VA) under ideal conditions. In our study, mean accuracy over all recordings was M = 1.73°, SD = 0.53°. Accuracy was not significantly different between groups ($t(222.32) = -1.76$, $p = 0.08$).

Data Preprocessing and Face Detection

To preprocess the data, we extracted fixations using the fixation detection algorithm provided by the Pupil Labs player software (Kassner et al., 2014). Fixations were detected using a maximum dispersion threshold of 1.5° VA, with durations ranging from 80 ms to 220 ms. Blinks (short drops in pupil detection confidence) were excluded during preprocessing using the blink detector plugin. Furthermore, all fixations below a confidence threshold of 0.8 were removed.

To locate facial AOIs, face detection was performed using a multi-task convolutional neural network (CNN) with the OpenCV and the facenet-pytorch package in python. This process identified the position of the experimenter's face and facial landmarks such as the eyes and corners of the mouth in each frame. Ellipses were fitted to the face, eyes, and mouth based on the bounding box returned by the face detection model. The size and shape of these ellipses adjusted dynamically to the orientation of the head and detected facial landmarks. Fixations where then mapped to AOIs frame by frame using a self-written python script. Lastly, fixations for each region were counted and averaged over time intervals (e.g. the total interview duration) resulting in a measure of the average fixation rate (fixations per second) on the respective region. This measure served as the dependent variable for further analysis. The python scripts used for face detection used in this study are publicly available in an online repository (https://github.com/raimund-buehler/face_detection_small) and described in more detail in our previous work (Buehler et al., 2024).

Pharmacological procedure

The pharmacological procedure consisted of 4 medication conditions (NAL/OXT, NAL/PLA, OXT/PLA, and PLA/PLA), each of which was experienced by each participant on 4 different testing days (within-subject). The order of the conditions was randomized among participants. Drugs were supplied by the General Hospital of Vienna (AKH). Naltrexone (Dependex, 50 mg) was administered orally in pill form. Oxytocin (Syntocinon,

5IU/ml) was administered via a nebulizer (Pari 128 G1000 Sinus 2, Pari, Germany). The solution was applied to both nostrils for an interval of 5 minutes each, resulting in an applied dosage of 10IU or 2ml solution.

This dose, which is lower than the conventional 24IU often encountered in oxytocin studies, was chosen for several reasons: First, previous research, mostly by Quintana and colleagues, has suggested that lower doses of oxytocin are more effective in eliciting behavioral effects of social cognition and attention, compared to higher doses (Martins et al., 2022; Quintana et al., 2015, 2016). Another critical consideration is the interactive effect of naltrexone on oxytocin, which increases the potential for cross-activation of vasopressin receptors that occur at higher doses of oxytocin (Neumann & Landgraf, 2012). Vasopressin receptors have been implicated in social behavior (Neumann & Landgraf, 2012; Quintana et al., 2021; Rilling et al., 2012) and their activation could confound the behavioral effects of oxytocin. Thus, the lower dose was chosen due to the effectiveness of lower doses suggested by recent literature and the goal of avoiding cross-activation of vasopressin receptors under simultaneous administration of naltrexone.

A pill with mannitol filler and a saline solution (0.9% NaCL) was administered in placebo conditions via the same rout of administration (i.e. oral intake and nebulizer) of the active drug. Thus, participants always received a pill and nebulizer solution in each session. The duration until peak plasma levels was accounted for in the timing of drug administration (naltrexone was administered 50 minutes and oxytocin 20 minutes before the start of testing) (Gonzalez & Brogden, 1988; Gossen et al., 2012). To minimize potential carryover effects, we aimed for an interval of 2 weeks between sessions. However, due to study logistics, some variation between testing dates could not be avoided. The mean interval between session dates was $M = 15.8$ days, $SD = 7.44$ days.

At the end of each session, participants were asked to guess the type of medication that was received, separately for the pill and the nebulizer. Blinding was successful for the pill, as indicated by a non-significant chi-square test ($\chi^2(1) = 2.53$, $p = 0.112$). Regarding the nebulizer, the chi-square test was significant ($\chi^2(1) = 6.83$, $p = 0.009$), indicating that blinding was potentially compromised. Given that oxytocin tends to have minimal side effects, it is possible that the solution containing the medication was discernible by smell.

However, when examining nebulizer blinding for individual medication conditions (with and without naltrexone), the chi-square test was only significant for the PLA/OXT vs PLA/PLA comparison ($\chi^2(1) = 6.76$, $p = 0.009$), while the NAL/OXT vs NAL/PLA comparison was not significant ($\chi^2(1) = 0.96$, $p = 0.328$). If participants had been able to reliably discern the medication by smell, this effect should have reflected in a significant result across both comparisons. An alternative interpretation for the significant result is a response bias towards placebo rather than the ability to discern nebulizer solutions: In the PLA/OXT and PLA/PLA conditions, participants were most likely to correctly identify placebo (41 hits). Since oxytocin's effects are not consciously noticeable, participants may have guessed 'placebo' more often, inflating correct placebo responses and driving the significant chi-square test. Thus, while participants may have been able to correctly identify the nebulizer condition by smell, alternative explanations, such as a response bias towards placebo, are plausible.

Questionnaire Measures - AQ and SIAS

To measure autistic traits, a German shortened version of the Autism Spectrum Quotient (AQ-k; Freitag et al., 2007) was used. The AQ-k contains 33 items (e.g. "*I prefer to do things on my own rather than with others.*") and is suitable for adults and adolescents aged 16 years and above with normal intellectual functioning. For screening purposes of ASD in clinical practice, a cut-off of 17 was proposed (Freitag et al., 2007). Items were scored using a binary system, with scores of 0 or 1 for each item, resulting in a maximum score of 33. Higher scores indicate higher levels of autistic traits (Ruzich et al., 2015). Additionally, scores were computed for 3 subscales, *social interaction and spontaneity*, *imagination,* and, *communication and reciprocity* (Freitag et al., 2008).

Social anxiety was measured using the social interaction anxiety scale (Heimberg et al., 1992). This questionnaire consists of 20 items (e.g. *"I feel tense if I am alone with just one other person"* or *"I have difficulty making eye contact with others"*). Items were scored using a 5-point Likert-scale, with a maximum score of 80. Higher scores indicate higher levels of social anxiety. A score of 34 is commonly used as a cutoff suggesting social anxiety disorder (SAD) (Heimberg et al., 1992).

We focused on these two questionnaires in our analyses, as they were most relevant to our research objectives, while other questionnaire measures reported in Table 1 (BDI, STAI-t, IRI, PANAS) primarily served to compare groups as control variables. Sympathy ratings are described in more detail in the supplementary material.

Cortisol

To investigate medication effects, saliva samples for cortisol measurement were collected at 3 timepoints: First, at the beginning of the session (baseline; T1), second after administration of medications, shortly before the start of the experiment during the estimated peak of medication plasma levels (T2). Third, at the end of the experiment (T3). IBL SaliCaps (IBL International GmbH, Hamburg, Germany) were used to collect samples according to the manufacturer's protocol. Specifically, participants were asked to collect saliva in their mouths and refrain from swallowing for 2 minutes, which was stopped with a timer. Saliva was then transferred into the SaliCap container via a small plastic straw.

Samples were stored in a freezer at -30°C until analysis. All analyses were conducted by Dresden LabService (Dresden, Germany). After thawing, SaliCaps were centrifuged at 3,000 rpm for 5 min. Salivary concentrations were measured using commercially available chemiluminescence immunoassay with high sensitivity (IBL International, Hamburg, Germany). The intra and interassay coefficients for cortisol were below 8%. In ten percent of samples (79 samples), a dual estimation of cortisol levels was conducted, with an average deviation between estimations of $M = 3\%$.

Statistical analysis: Eye Tracking

To investigate the effects of group and medication, we used a linear mixed model (LMM) with a random intercept by subject. The dependent variable of interest was fixations rates, defined as the number of fixations per second averaged over a certain time interval, i.e. the total interview duration or the duration of each interview question. Predictors were specified with a three-way interaction in the following way: group (2 levels: ASD / CTRL) × AOI (4 levels: eyes / mouth / face / background) × medication (4 levels: NAL/OXT; PLA/OXT; NAL/PLA; PLA/PLA):

$$Fixations/Second \sim Group \times AOI \times Medication + (1 \mid Subject) \quad (M1)$$

In addition, we specified a model to examine differences between open and closed questions. As described above, questions were categorized as closed (1-6) or open (7-13) questions. To reduce model complexity, we focused exclusively on fixations on the eyes, removing AOI as a predictor. In this case, the dependent variable (fixation rates) was averaged over the duration of each question rather than the entire interview. The model retained the group and medication predictors while also including question type, for which an additional intercept nested within subject was specified:

$$Fixations\ on\ Eyes/Second \sim Question\ Type \times Group \times Medication + (1 \mid Subject / Question\ Type) \quad (M2)$$

To further explore the effects of autistic traits and social anxiety, the same model as M1 was run, but with AQ and SIAS scores replacing the group predictor.

$$Fixations/Second \sim AQ \times AOI \times Medication + (1 \mid Subject) \quad (M3)$$

$$Fixations/Second \sim SIAS \times AOI \times Medication + (1 \mid Subject) \quad (M4)$$

Lastly, to investigate whether social anxiety mediated the relationship between autistic traits and fixation rates, we conducted mediation analyses using SIAS scores as a mediator. Specifically, we tested two separate mediation models: one examining whether the effect of group (ASD vs. CTRL) on fixation rates was mediated by social anxiety (SIAS scores), and another testing whether AQ scores were mediated by social anxiety. In both cases, the mediation models were specified with fixation rates on the eyes as the dependent variable, the respective predictor (group or AQ), and SIAS scores as the mediator.

Statistical Analysis: Cortisol

Salivary cortisol levels were measured at three time points during each session: baseline (T1), pre-task (T2), and post-task (T3). Raw cortisol levels were log-transformed to meet model assumptions. Outliers exceeding 1.5 times the interquartile range were removed to ensure robust results (60 of 789 data points). The model included group, medication and timepoint as fixed effects, with a random intercept for both subject and session to account for repeated measures across sessions. The model for raw cortisol levels was specified as follows:

$$Log(Cortisol\ Level) \sim Group \times Medication \times Timepoint + (1 \mid Subject / Session) \quad (M5)$$

Post-hoc pairwise comparisons were conducted to further explore significant interactions. These comparisons assessed differences across medication conditions within each group, as well as changes in medication effects over timepoints. Plots were generated to visualize cortisol fluctuations over time for each medication condition and group.

As raw levels do not capture the full temporal dynamics of cortisol responses, the area under the curve (AUC) is typically used to quantify both magnitude and duration of cortisol secretion (Pruessner et al., 2003). Two common measures are the area under the curve with respect to ground (AUCg), which reflects total cortisol output including baseline levels, and the area under the curve with respect to increase (AUCi), which aims to capture cortisol reactivity by accounting for baseline differences. We were mainly interested in cortisol reactivity (AUCi). This measure was computed by subtracting the first timepoint (T1) from all subsequent values before calculating the area under the curve. A linear mixed model was then used to analyze AUCi, with group and medication as fixed effects and a random intercept for participant:

$$AUCi \sim Group \times Medication + (1 \mid Subject) \qquad (M6)$$

Although AUCi provides a baseline-corrected measure of cortisol reactivity, it remains dependent on the initial baseline measurement, which may have been influenced by individual differences upon arrival at the lab. Furthermore, sessions were conducted at different times of day, which could confound baseline measures due to diurnal fluctuations in cortisol. To address these potential confounds, we also used an approach akin to the MaxMin method (Miller et al., 2018) to obtain a measure of cortisol reactivity independent of baseline levels. For this, we eliminated baseline measures from calculations and computed the difference between timepoint 2 and timepoint 3 (T2T3), providing a baseline-free estimate of cortisol reactivity. This measure proved to capture the effects of medications on cortisol more robustly (see results section). A linear mixed model was then used to analyze T2T3, with group and medication as fixed effects and a random intercept for participant, equivalent to M6:

$$T2T3 \sim Group \times Medication + (1 \mid Subject) \qquad (M7)$$

Lastly, to examine the relationship between cortisol reactivity and social attention, we used AUCi and T2T3 as predictors of fixation rates on the eye region. Separate linear mixed models were specified for each measure, with group, medication, the respective cortisol reactivity metric (AUCi or T2T3) as fixed effects, and a random intercept for participant:

$$Fixations\ on\ Eyes/Second \sim AUCi \times Group \times Medication + (1 \mid Subject) \qquad (M8)$$

$$Fixations\ on\ Eyes/Second \sim T2T3 \times Group \times Medication + (1 \mid Subject) \qquad (M9)$$

Control Analysis: Menstrual Cycle

Fluctuations in oxytocin and sex hormones (such as estrogen and progesterone) during the menstrual cycle can significantly influence responses to oxytocin (Engel et al., 2019). To control for these, we recorded the date of last menstruation (first day of menstrual bleeding), as well as the length of the menstrual cycle for female participants during prescreening. From this information, menstrual phases on testing dates were computed for each participant. Phases were defined as follows: Menstrual (day 1 - 5), Follicular (day 6 - 13), Ovulatory (day 14 - 16), Luteal (day 17 and above). Data was missing for 7 female participants, who failed to provide the information in the questionnaire. 4 participants were tested more than 120 days after providing the date and were excluded from computation, due to accumulating error. Data for menstrual phases was computed for 28 participants. Of these, 26 did not use any kind of hormonal contraception and 2 used a vaginal ring for contraception.

To investigate differences in fixations on the eyes between cycle phases, we specified a model with fixations on the eyes as the dependent variable and menstrual cycle phase (4 levels: Menstrual / Follicular / Ovulatory / Luteal) as an additional interacting predictor. These results are reported in the supplementary material:

$$Fixations\ on\ Eyes/Second \sim Cycle\ Phase \times Group \times Medication + (1 \mid Subject) \qquad (SM1)$$

Software and Final Remarks

For all models involving fixation rates, a square root transformation was used to meet model assumptions. Session number and self-reported participant gender were added as control variables and fixed effects for all models, but are not included in the pseudo-code above for simplicity. Analyses on the percentage of total fixation duration, an alternative dependent variable, are reported in the supplementary material. All statistical analyses were performed using R 4.2.2 (R Core Team, 2022). LMM's were fit via maximum likelihood using the *lmerTest* package version 1.1.31. F-tests were conducted using the *anova()* command, which computes degrees of freedom for linear mixed models using Satterthwaite's method (Kuznetsova et al., 2017). Contrasts and t-tests were conducted using the *emmeans* package with the default Kenward-Roger method for the computation of degrees of freedom and the default Tukey method for correction of multiple comparisons (Lenth, 2022). Plots were created using the *ggplot2* package (Wickham, 2016).

3 Results

Eye Tracking – Group and Medication Differences (M1)

F-tests for the M1 model, which included interactive predictors of AOI, group, and medication, revealed a significant main effect of AOI ($F(3, 892.24) = 123.17$, $p < 0.0001$). Participants made most fixations on the background ($M = 1.12$, $SD = 0.32$) and eye region ($M = 0.96$, $SD = 0.42$), with significantly fewer fixations on the mouth ($M = 0.61$, $SD = 0.38$) and face regions ($M = 0.72$, $SD = 0.23$). Specifically, participants made more fixations on the eyes compared to the mouth ($t(890) = 11.73$, $p < 0.0001$) and face ($t(890) = 8.01$, $p < 0.0001$) and more on the background compared to the eyes ($t(890) = -5.69$, $p < 0.0001$), mouth ($t(890) = -17.42$, $p < 0.0001$), and face ($t(890) = -13.70$, $p < 0.0001$).

No other significant main effects were observed. However, the interaction between group and AOI was significant ($F(3, 892.24) = 5.20$, $p = 0.001$). Post-hoc contrasts indicated that groups were significantly different for fixations on the eye region ($t(285) = -2.04$, $p = 0.043$, Cohen's $d = -0.24$, 95% CI [-0.47, -0.01]) and the background region ($t(285) = 2.29$, $p = 0.023$, Cohen's $d = 0.27$, 95% CI [0.04, 0.50]) when collapsing over medications. This suggests that CTRL participants made more fixations on the eye region and fewer fixations on the background compared to ASD participants conditions (see Figure 2).

Regarding the effect of medications, the variance between medications was small, and no significant interactions with other predictors were observed (all $p > 0.05$). This suggests that medications did not significantly influence fixation rates. However, further analysis of between-group contrasts revealed that the group difference for fixations on the eye region was significant only in the PLA/PLA condition ($t(873) = -2.03$, $p = 0.042$), with ASD participants exhibiting fewer fixations on the eyes compared to CTRL participants.

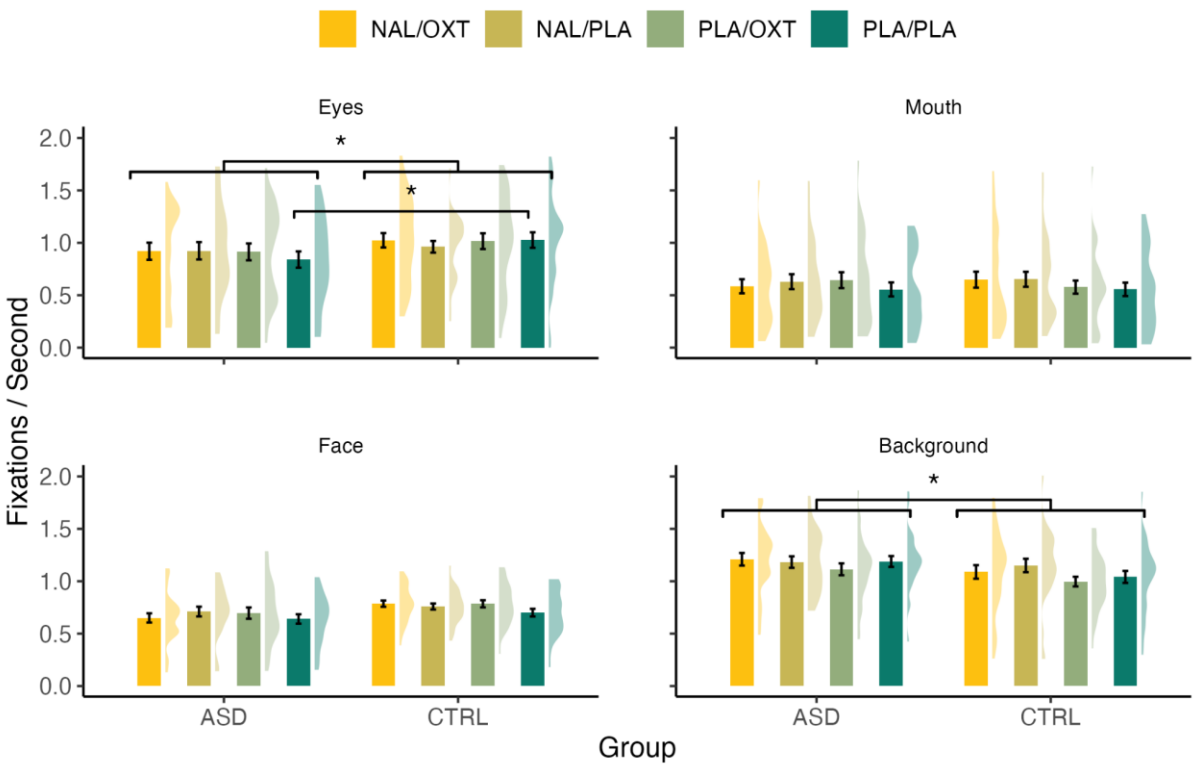

Figure 2 (preliminary) - Fixations per second on different AOIs across groups and medications. Group differences are significant for the eye and background regions when collapsing over medications. Groups are also significantly different for the PLA/PLA condition for fixations on the eye region in post-hoc contrasts comparing individual medication conditions. * $p < 0.05$

Timecourse Analysis (M2)

The unique group difference in the placebo condition became even more evident when fixation rates on the eyes were averaged over smaller time windows, specifically the duration of each interview question (model M2; see Figure 3). Model M2 retained group and medication as predictors but focused exclusively on fixations on the eyes, removing the AOI predictor. It also included question type (open vs. closed) as an additional predictor. F-tests revealed a significant main effect of question type (open vs. closed) ($F(1, 55.58) = 8.64$, $p = 0.0048$), as well as a significant interaction between group and medication ($F(3, 2894.24) = 7.28$, $p < 0.001$). Furthermore, a trend was observed for the interaction between group and question type ($F(1, 55.58) = 3.56$, $p = 0.064$).

Participants made significantly more fixations on the eyes during closed questions compared to open questions when collapsing across groups and medications ($t(66.7) = 2.94$, $p = 0.0046$). Examining the group by medication interaction revealed that groups differed significantly in the PLA/PLA condition ($t(76.5) = -2.26$, $p = 0.027$), but not in other conditions (all $p > 0.05$), consistent with findings from model M1 reported above.

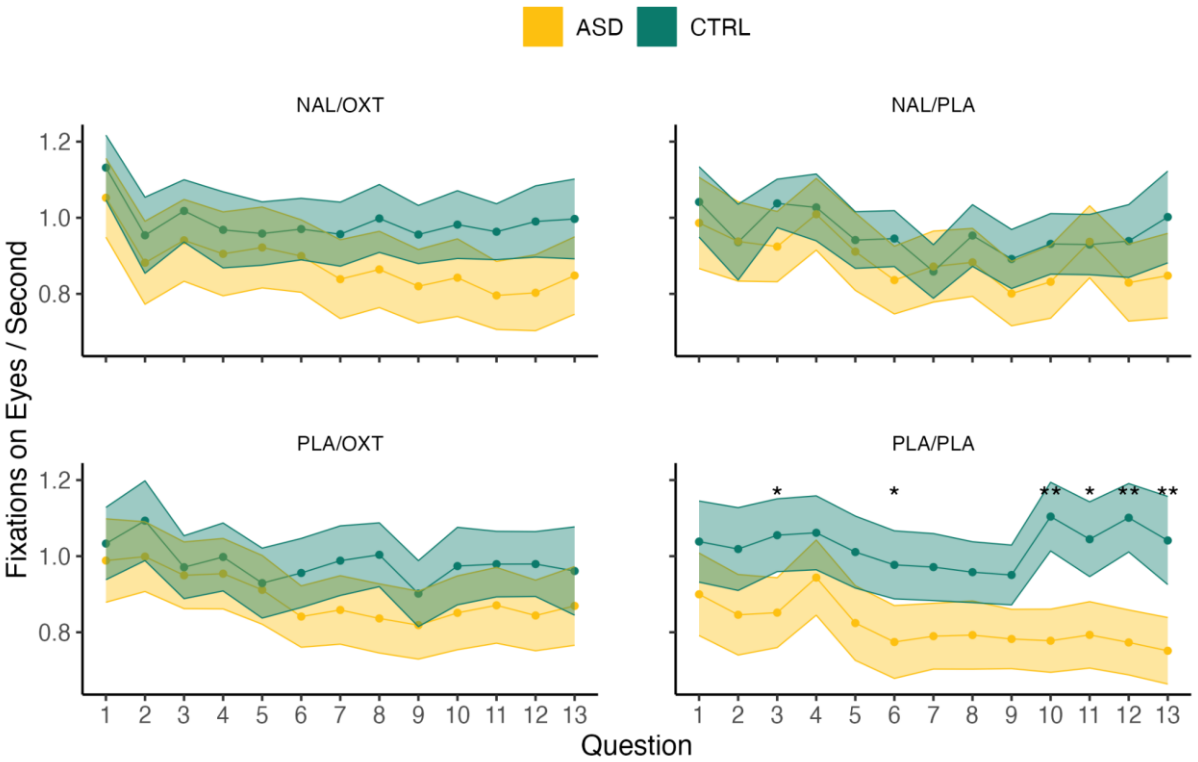

Figure 3 (preliminary) - Fixations on the eye region, averaged over the duration of interview questions (1-13). Stars mark significant group differences. Questions 1-6 were categorized as closed questions, which could be answered by a simple "yes" or "no". Questions 7-13 were categorized as open questions, potentially demanding more social engagement. A list of all questions can be found in the supplementary material. * $p < 0.05$, ** $p < 0.01$

As illustrated in Figure 3, group differences were significant for some questions in the PLA/PLA condition but not in other medication conditions. At first glance, this seems to be driven by an increase in fixations on the eyes for the control group towards the end of the interview. However, closer analysis of the group by question type trend, comparing open vs. closed questions (or later vs. earlier parts of the interview), indicated that ASD participants generally fixated less on the eyes during later parts of the interview ($t(64.3) = 3.33$, $p = 0.0015$). No difference between earlier and later parts of the interview were observed in the control group ($t(69.4) = 0.76$, $p = 0.448$), suggesting that gaze behavior of the CTRL group remained constant throughout the interview.

When taking medication into account, the group difference in the PLA/PLA condition also appeared to be driven by the ASD group, which made fewer fixations on the eyes for later parts of the interview, ($t(314) = 2.05$, $p = 0.041$). In contrast, this comparison was not significant for the control group, ($t(350) = -0.42$, $p = 0.676$). Nonetheless, these findings might reflect random fluctuations in eye gaze and should be interpreted with caution. We explore this observation in more detail in the discussion section.

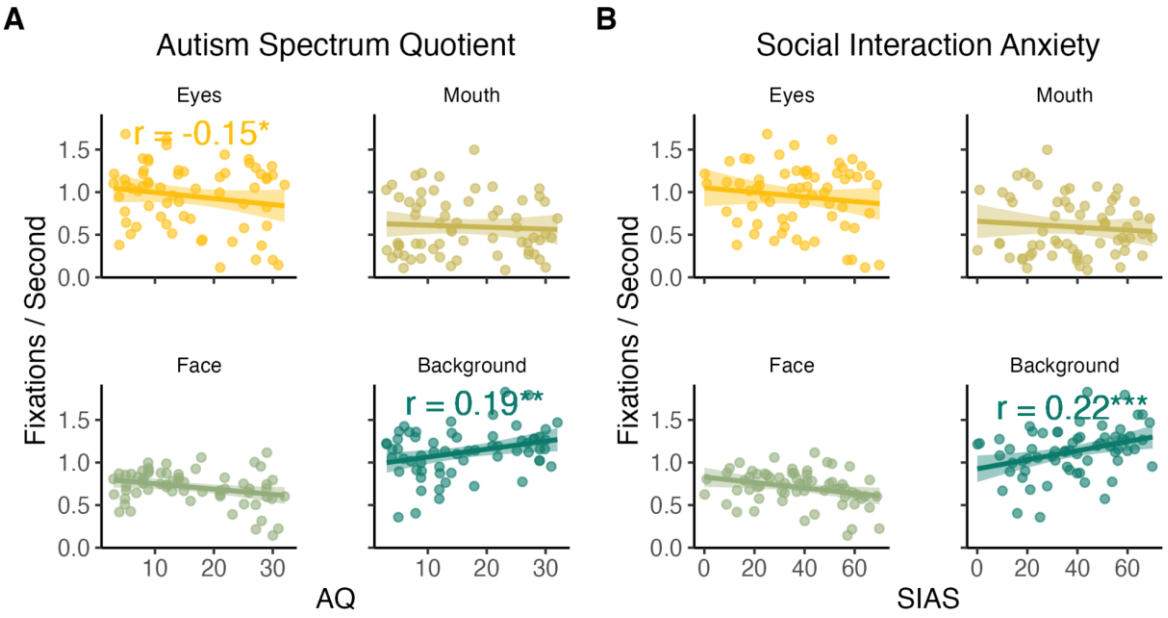

Figure 4 (preliminary) - A Relationship of fixation rates on different AOI's and autistic traits (AQ, left) and **B** social interaction anxiety (SIAS, right). Data was collapsed over groups and medication conditions. * $p < 0.05$

AQ and SIAS (M3 and M4)

Replacing the group predictor with AQ and SIAS scores revealed a very similar pattern of results (Figure 4). For AQ, a significant main effect of AOI ($F(3, 852.28) = 118.52$, $p < 0.0001$) and a significant interaction of AQ × AOI ($F(3, 852.28) = 8.00$, $p < 0.0001$) were observed. Post-hoc contrasts suggested a significant negative effect of AQ for the eye region ($t(234) = -2.27$, $p = 0.024$, $r = -0.15$, 95% CI [-0.27, -0.02]) and a positive effect for the background region ($t(234) = 2.95$, $p = 0.004$, $r = 0.19$, 95% CI [0.06, 0.31]).

Similarly, we observed a significant interaction of SIAS × AOI ($F(3, 859.67) = 8.18$, $p < 0.0001$), with a significant effect for the background region ($t(235) = 3.44$, $p < 0.001$, $r = 0.22$, 95% CI [0.09, 0.33]). However, SIAS was not a significant predictor for fixations on the eye region ($t(235) = -1.39$, $p = 0.166$). Due to the strong association of SIAS with ASD and AQ scores (see Figure 5), we further investigated whether the effects of AQ and group were mediated by social anxiety scores.

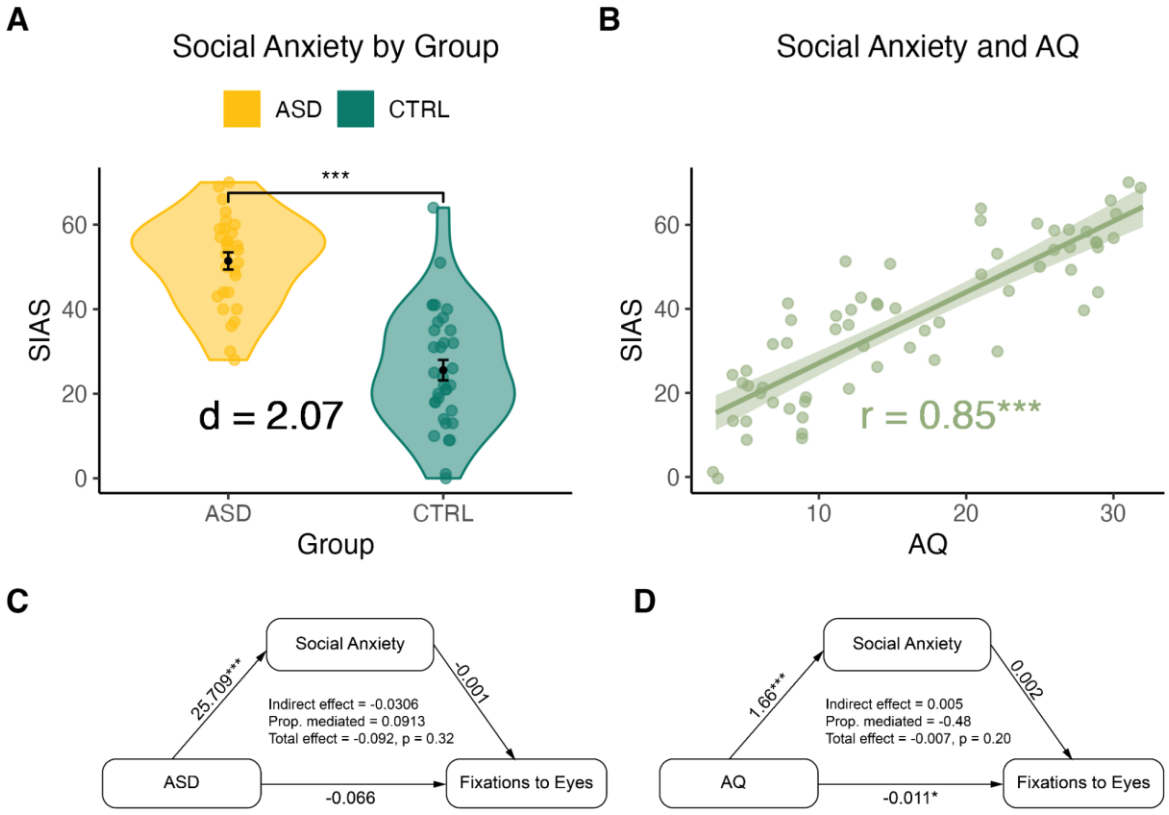

Figure 5 - **A** Group differences in social interaction anxiety scale (SIAS) scores. **B** Correlation of autistic traits (AQ) and SIAS scores. **C** Mediation model of the effect of ASD on fixations on the eye region via SIAS scores. **D** Mediation model of the effect of AQ on fixations on the eye region via SIAS scores. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Mediation Analyses

To test this hypothesis, we conducted a mediation analysis examining both AQ and group effects via SIAS. While both group and AQ had a strong impact on SIAS scores (Cohen's d = 2.07, 95% CI [1.75, 2.38], $p < 0.0001$; r = 0.85, 95% CI [0.85, 0.88], $p < 0.0001$), neither predictor showed significant mediation through SIAS (see Figure 5C and D for full results). Moreover, when SIAS was included as a predictor, the group difference was no longer significant ($p = 0.40$), likely reflecting a suppression effect due to multicollinearity, given the high correlation between the predictors. As a result, this finding should be interpreted as a statistical artifact rather than evidence of a meaningful relationship. The same pattern emerged in the mediation analysis with AQ scores; however, in this case, the negative effect of AQ on fixations on the eye region remained significant ($t(234) = -1.975$, $p = 0.0494$).

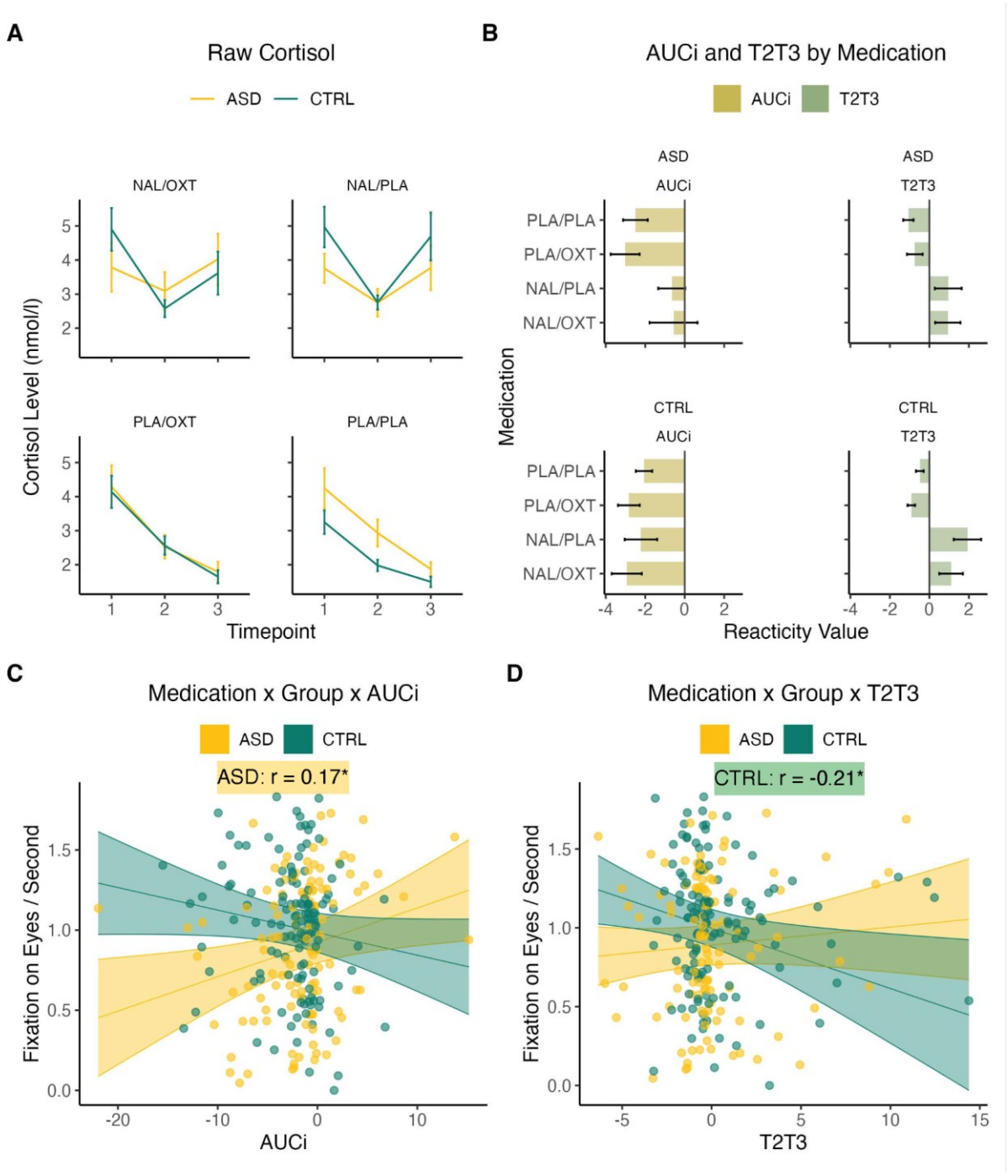

Figure 6 (preliminary) - **A** Raw cortisol levels across timepoints. A cortisol response can be observed in sessions involving NAL between T2 and T3, other conditions are marked by general decline of cortisol levels . **B** AUCi (left) and T2T3 (right): NAL conditions generally elicit a cortisol response compared to placebo conditions. This is more clearly captured by T2T3, where differences between NAL and other conditions are all significant when collapsing over groups (all $p < 0.01$; not shown) **C** Interaction of AUCi and fixations on eyes. AUCi is positively associated with fixations on the eyes in ASD when collapsing over medications **D** Interaction of T2T3 and fixations on eyes. T2T3 is negatively associated with fixations on the eyes in the CTRL group, particularly in the PLA/PLA condition. * $p < 0.05$.

Cortisol – Raw Cortisol Levels (M5)

The analysis of raw cortisol levels revealed significant main effects of timepoint ($F(2, 635) = 49.06$, $p < 0.001$) and medication ($F(3, 657) = 14.36$, $p < 0.001$). No significant main effect of group was observed ($p = 0.377$). However, the interactions of group × medication ($F(3, 630) = 3.65$, $p = 0.012$) and medication × timepoint ($F(6, 635) = 7.35$, $p < 0.001$) were significant.

Contrasts revealed a decrease in cortisol levels across timepoints, with higher cortisol levels for earlier timepoints compared to later ones when collapsing over medications and groups: the comparison between T1 and T2 was significant ($t(641) = 7.67$, $p < 0.001$), as was the comparison between T1 and T3 ($t(641) = 9.25$, $p < 0.001$).

Additionally, sessions involving NAL administration (NAL/OXT and NAL/PLA) differed from those without NAL when collapsing over timepoints and groups. Specifically, significantly higher cortisol levels were observed in the NAL/PLA condition compared to both PLA/OXT ($t(661) = 5.6$, $p < 0.001$) and PLA/PLA ($t(650) = 5$, $p < 0.001$). Similarly, cortisol levels were higher in the NAL/OXT condition compared to PLA/OXT ($t(669) = 3.87$, $p < 0.001$) and PLA/PLA ($t(655) = 3.27$, $p = 0.006$).

Examining timepoints in sessions involving NAL more closely, cortisol levels first fell between T1 - T2 and rose between T2 - T3 (see Figure 6, A). Accordingly, probing the timepoint × medication interaction suggested higher cortisol levels for both NAL/PLA as well as NAL/OXT at T3 compared to other medications (all p-values < 0.001), but not at other timepoints. This indicates that the effect of medication on cortisol was captured most effectively by the difference between T2 and T3.

Medication differences seemed to be driven by the CTRL group when collapsing over timepoints. Specifically, for the CTRL group, NAL/OXT differed from both PLA/OXT ($t(644) = 3.2$, $p = 0.008$) and PLA/PLA ($t(555) = 4.21$, $p < 0.001$) and NAL/PLA differed from both PLA/OXT ($t(660) = 4.82$, $p < 0.001$) and PLA/PLA ($t(653) = 5.8$, $p < 0.001$). In contrast, in the ASD group, only the comparison between NAL/PLA and PLA/OXT was significant ($t(662) = 3.15$, $p = 0.009$), with no other medication contrasts reaching significance. However, this may have been confounded by baseline differences, which can obscure group differences in medication effects when collapsing over timepoints. For this reason, we analyzed two measures of cortisol reactivity with the aim to account for baseline differences: the area under the curve with respect to increase (AUCi) and the difference between T2 and T3 (T2T3).

AUCi (M6)

The analysis of the area under the curve with respect to increase (AUCi) suggested a decrease of cortisol over timepoints consistent with the analysis of raw cortisol levels, as means were negative over all conditions and groups (see Figure 6, B). A trend was observed for a main effect of medication ($F(3, 173) = 2.29$, $p = 0.080$), but not for group ($p = 0.187$). Post-hoc tests suggested a marginal difference between NAL/PLA and PLA/OXT when collapsing over groups ($t(174) = 2.34$, $p = 0.094$). No other contrast for medication was significant when collapsing over groups.

A near significant trend was also observed for the group × medication interaction ($F(3, 174) = 2.63$, $p = 0.052$). Unraveling this interaction suggested that NAL was mainly effective in the ASD group. Specifically, within the ASD group, both NAL/PLA and NAL/OXT significantly differed from PLA/OXT ($t(174) = 2.62$, $p = 0.047$); ($t(175) = 2.92$, $p = 0.021$), suggesting a cortisol response for sessions in which NAL was given. In contrast, within

the CTRL group, no contrasts between medications were significant, however this may again be confounded by high baseline levels of cortisol.

AUCi and Fixations on Eyes (M8)

When investigating the relationship of AUCi and fixations on the eye region, the interaction between group × AUCi was significant ($F(1, 202) = 9.34$, $p = 0.003$). Post-hoc tests suggested a positive relationship between AUCi and fixations on the eyes in the ASD group when collapsing over medications ($t(202) = 2.41$, $p = 0.017$, $r = 0.17$). In the CTRL group, the effect of AUCi was negative and approached significance ($t(182) = -1.79$, $p = 0.076$). No other main effects or interactions were observed. Thus, increases in cortisol seem to be associated with more fixations on the eye region in the ASD group, but not the CTRL group. However, as AUCi may be influenced by baseline measures, we also investigated the relationship of a baseline-free measure of cortisol reactivity, T2T3, and its relationship with fixations on the eye region.

T2T3 (M7)

Consistent with the raw cortisol analysis, a significant main effect of medication was observed ($F(3, 185.03) = 11.65$, $p < 0.001$), while no significant main effect of group ($p = 0.269$) or interaction of group × medication ($p = 0.672$) was found. Post-hoc pairwise comparisons revealed that cortisol reactivity varied across medication conditions when collapsing over groups. Specifically, sessions involving NAL administration seemed to clearly elicit a cortisol response when compared to sessions without NAL. For example, NAL/PLA was associated with significantly higher T2T3 values compared to both PLA/OXT ($t(181) = 4.60$, $p < 0.001$) and PLA/PLA ($t(178) = 4.45$, $p < 0.001$). Similarly, NAL/OXT yielded significantly greater cortisol reactivity than PLA/OXT ($t(180) = 3.84$, $p = 0.001$) and PLA/PLA ($t(180) = 3.69$, $p = 0.002$). However, no significant difference was found between NAL/OXT and NAL/PLA ($p = 0.826$), suggesting that both NAL conditions elicited similar cortisol reactivity increases relative to non-NAL conditions.

Unlike in the analysis of AUCi, no trend of an interaction between group and medication was observed, suggesting that medication effects on cortisol reactivity, as captured by T2T3, were not modulated by group. This may reflect a more consistent effect of NAL on cortisol changes between T2 and T3, independent of baseline variations, highlighting the utility of T2T3 as a baseline-free measure of cortisol reactivity.

T2T3 and Fixations on Eyes (M9)

To examine the relationship between cortisol reactivity and social attention, we analyzed the interaction between T2T3 and fixation rates to the eye region. The analysis revealed a significant interaction between group and T2T3 ($F(1, 174.16) = 6.19$, $p = 0.014$), suggesting that cortisol reactivity as captured by T2T3 influenced social attention differently between groups. No significant main effects of medication ($p = 0.731$) or group ($p = 0.228$) were observed, nor a three-way interaction between group, medication, and T2T3 ($p = 0.082$).

Post-hoc trend analysis of the interaction between group and T2T3 revealed that in the ASD group, T2T3 was not significantly associated with fixation rates ($p = 0.357$). However, in the CTRL group, a significant negative association was observed ($t(171) = -2.42$, $p = 0.016$, $r = -0.21$), indicating that higher cortisol reactivity was associated with reduced fixation rates on the eyes in the CTRL group.

Further examining the group × medication × T2T3 interaction indicated that the observed effect in the CTRL group was mainly driven by the PLA/PLA condition. Specifically, in the CTRL group, a significant negative association between T2T3 and fixation rates was found only in the PLA/PLA condition ($t(172) = -2.78$, $p = 0.006$), whereas no such relationship was observed in other medication conditions. In contrast, no significant associations were found in the ASD group across any medication conditions. These findings suggest that in the CTRL group, heightened cortisol reactivity as captured by T2T3 is associated with reduced social attention, particularly in the absence of pharmacological manipulation.

Cortisol Analysis: Summary

In summary, a prominent cortisol response was observed in the NAL conditions across both groups, reflected in both AUCi and T2T3, with T2T3 capturing the effect most clearly. In the ASD group, AUCi was positively associated with fixations on the eyes, suggesting that higher cortisol reactivity corresponded to increased social attention. In contrast, in the CTRL group, higher T2T3 values were associated with fewer fixations on the eyes, indicating that greater cortisol reactivity may have led to reduced social attention.

This suggests that group differences in social attention may have been influenced by cortisol reactivity, which was modulated by naltrexone intake. This could explain why the significant difference between groups on fixations on groups observed in the placebo condition disappeared under naltrexone administration. However, given the exploratory nature of these results, further research is needed to establish a clear relationship between cortisol, naltrexone and fixations on eyes.

4 Discussion

Atypical gaze behavior, particularly reduced fixations on the eye region, is a well-documented finding in ASD (Chita-Tegmark, 2016) and has been associated with both opioids and oxytocin (Fan et al., 2020). In this study, we investigated the interactive effects of oxytocin and naltrexone on social attention during a naturalistic face-to-face interaction in neurotypical and ASD participants. Additionally, given the association of HPA axis activity and social behavior (Sandi & Haller, 2015), we explored the role of cortisol reactivity as a potential modulator of social gaze across groups. As prior research has demonstrated that oxytocin and naltrexone can interactively affect social attention in non-human primates (Dal Monte et al., 2017), we anticipated a stronger effect on eye gaze in the combined condition than with either pharmacological agent alone.

Across our analyses, we observed a group difference in fixations on the eye region, with ASD participants fixating less often on the eyes compared to neurotypical controls. This finding aligns with previous research conducted in screen-based settings (Chevallier et al., 2015; Chita-Tegmark, 2016; Frazier et al., 2017) and is consistent with clinical observations of reduced eye contact in ASD (Senju & Johnson, 2009). In contrast to our expectations, differences between medication conditions were not statistically significant. However, post-hoc analyses revealed that group differences were most pronounced in the placebo condition and appeared attenuated under individual and combined administration of naltrexone and oxytocin. This suggests that both medications may have diminished baseline group differences.

Several factors may have contributed to the observed results. The broader literature suggests that oxytocin's behavioral effects are highly context-dependent (Quintana, 2018), and our findings align with this perspective. While both oxytocin and naltrexone have well-documented effects in screen-based experiments, their impact on social attention in naturalistic settings may be more subtle, potentially attenuating group differences rather than driving within-group differences. Nonetheless, caution is warranted in interpreting this result given the lack of a significant drug-by-group interaction.

A similar pattern emerged when categorical group comparisons were replaced with continuous measures of autistic traits and social interaction anxiety. Specifically, autistic traits were related to reduced fixations on the eyes, while social anxiety was associated with increased fixations on the background region. This supports the view that social attention varies along a continuum and implicates potential mediating variables linking autistic traits to eye gaze behavior. However, although both ASD diagnosis and autistic traits were strong predictors of social interaction anxiety, mediation analyses revealed that neither the group effect nor the effect of autistic traits was significantly mediated by social anxiety. This finding may suggest that reduced social attention in ASD reflects hypo-arousal induced indifference towards the social environment rather than active avoidance of eye contact (Moriuchi et al., 2017).

The analysis of cortisol reactivity provided additional insights into stress and arousal-related mechanisms potentially underlying group differences. As expected, naltrexone elicited a robust cortisol response both in isolation and in the combined condition, consistent with HPA axis disinhibition under μ-opioid antagonism (King, 2002). This suggests that naltrexone exerted a dominant effect in the combined condition. In contrast, oxytocin did not significantly affect cortisol levels, neither when administered alone nor in combination. Previous findings suggest that oxytocin's effects on the HPA axis are most pronounced under acute stress (Cardoso et al., 2014). The lack of oxytocin's effect on cortisol levels in our study may thus reflect the absence of a stress induction paradigm. Additionally, dosage is a critical factor influencing oxytocin's effect on cortisol levels (Cardoso et al., 2013). The

dosage used in our study may have been too low to elicit measurable effects on cortisol levels, a limitation which is discussed in more detail below.

Importantly, group differences emerged when examining the relationship between cortisol reactivity and eye-tracking. Specifically, cortisol reactivity was positively associated with fixations on the eyes in the ASD group ($r = 0.17$), whereas a negative association was observed in the control group ($r = –0.21$). This leaves room for the interpretation that attenuated group differences in eye gaze under opioid antagonism resulted from arousal-related mechanisms, which acted differently on gaze behavior across groups. Recent models emphasize the balance between opioid activity and the social avoidance system (SAS), which encompasses HPA axis activity (Pellissier et al., 2018). Additionally, theories of hypo- and hyperarousal in ASD suggest that arousal-related mechanisms may differentially influence gaze behavior across groups (Cuve et al., 2018).

Previous research in healthy adults has demonstrated that individuals exhibiting stronger cortisol responses to acute social stress tended to avoid eye gaze (Vatheuer et al., 2021). Similarly, in our study, individuals in the control group with higher cortisol reactivity fixated less on the eyes. This could reflect a maladaptive stress response, which reduced the propensity to engage in direct eye gaze, potentially indicating hyperarousal-induced gaze avoidance. Conversely, those with lower cortisol responses may have exhibited a more optimal level of arousal, facilitating social attention. In contrast, in the ASD group, individuals with lower cortisol reactivity may have been in a state of hypo-arousal induced indifference, corresponding to the baseline reduction of eye contact often observed in ASD. Increased social gaze in ASD participants with higher cortisol reactivity may thus reflect more optimal levels of arousal, facilitating social attention. This would suggest that moderate activity of the HPA axis is indeed adaptive and facilitates social attention in ASD.

The differential association between cortisol reactivity and eye gaze in ASD versus control participants may also implicate the κ-opioid receptor (KOR) system. KORs, primarily activated by dynorphins, are known to modulate stress responses through their interaction with the HPA axis (Bruchas et al., 2010; Knoll et al., 2011). In rodent models, KOR antagonism facilitates social interaction, particularly following stress exposure (Lalanne et al., 2017; McLaughlin et al., 2006). In humans, naltrexone has been shown to increase attention to emotional facial expressions, an effect potentially mediated by KOR antagonism (Wardle et al., 2016). In our study, KOR antagonism may have contributed to adaptive arousal and increased social attention in the ASD group, whereas in controls, MOR antagonism may have induced a state of maladaptive hyperarousal and reduced gaze on the eye region. While speculative, these findings raise the possibility that naltrexone differentially modulates stress and attention across groups, potentially through divergent engagement of opioid receptor subtypes.

Lastly, like naltrexone, oxytocin also seems to have attenuated group differences, as no significant group difference was observed under oxytocin administration. However, unlike naltrexone, oxytocin did not elicit measurable changes in cortisol levels in our study, suggesting that its effects on social attention are unlikely to be mediated by stress-buffering mechanisms. Instead, oxytocin may have selectively enhanced social motivation or salience in the ASD group, consistent with previous research reporting increased attention to the eye region following oxytocin administration in individuals with ASD (Auyeung et al., 2015). Such an interpretation would be in line with models proposing context- and trait-dependent effects of oxytocin on social attention.

In summary, group differences in social attention could have emerged as a function of cortisol-related stress- and arousal, which were modulated by naltrexone. Naltrexone, via increased cortisol reactivity, may differentially influence social attention in ASD and CTRL groups, potentially attenuating group differences in

gaze behavior when no stress induction paradigm is administered. However, further research is needed to clarify the mechanisms underlying these interactions, including the distinct contributions of μ- and κ-opioid receptor systems and interindividual differences in stress responsivity. Pharmacological studies incorporating stress induction paradigms may help to disentangle these effects and help characterize how opioid and oxytocin modulation shapes social attention in ASD and neurotypical populations.

Limitations

Some limitations of the study also need to be taken into account. First, the oxytocin dosage of 10 IU could have been too low: An interactive effect of oxytocin and naltrexone on social attention was demonstrated using 24 IU in macaque monkeys (Dal Monte et al., 2017). While some studies suggest low doses of 8 IU oxytocin to be effective in eliciting effects at the neural level (Quintana et al., 2016), 10 IU may not have been the optimal dosage for our design. Future dose-response studies could help to determine an efficient dosage in naturalistic eye tracking studies. Second, drugs were administered via different routes: oxytocin was delivered nasally via a nebulizer, while naltrexone was delivered orally in pill form. Interactive effects of both drugs should occur in the central nervous system regardless of administration route. However, despite efforts to account for differences in uptake and time to reach peak plasma concentrations, these factors may have contributed to missing medication effects.

Additionally, there is substantial variability in individual sensitivity to both oxytocin and naltrexone, influenced by factors such as baseline oxytocin levels, opioid receptor density, and genetic polymorphisms in the oxytocin receptor gene (OXTR) or μ-opioid receptor gene (OPRM1) (Kumsta & Heinrichs, 2013; Troisi et al., 2011). Future studies could include genetic analyses to account for some of these factors. Furthermore, comorbid conditions and medication use in the ASD group could have interacted with or masked pharmacological effects.

Lastly, this study focused on investigating gaze behavior and trait measures of participants. However, social gaze is inherently bi-directional (Cañigueral & Hamilton, 2019). This problem is akin to the double empathy problem, which suggests that difficulties in social interactions arise not solely from the individual with ASD but from a mismatch between both interaction partners (Milton, 2012). To get a full picture of the interaction, both partners would need to be examined. Some studies have already employed dual eye tracking approaches to examine mutual gaze behavior of both interaction partners (Macdonald & Tatler, 2018; Mayrand et al., 2023; Rogers et al., 2018). A promising direction for future research is the application of non-invasive methods to measure gaze behavior in both participants while also capturing trait measures of the interaction partner, such as AQ or SIAS scores (Vehlen et al., 2021). Ensuring non-invasiveness is particularly important, as the presence of a head-worn eye tracker on the interaction partner may alter natural gaze behavior (Cañigueral et al., 2018).

Conclusions

To conclude, our results demonstrate both ASD and autistic traits consistently relate to atypical gaze patterns in naturalistic social interactions. This extends the literature on social attention, which has so far largely relied on screen-based tasks. The reduction of group differences in medication conditions aligns with a growing body of research suggesting that the behavioral effects of single doses of oxytocin and naltrexone in naturalistic social contexts may be more subtle than in controlled experimental settings. This raises the possibility that group differences in naturalistic settings may be attenuated under pharmacological modulation, rather than eliminated or reversed. Moreover, it suggests that effects observed in controlled paradigms may not fully translate to dynamic, realistic interactions.

At the same time, this study provides initial evidence that gaze behavior may be shaped by a complex interplay of social reward processing mediated by opioids and oxytocin as well as stress- and arousal-related mechanisms, which may not be fully captured in traditional screen-based paradigms. The observed effects of cortisol on social gaze further highlight the need to consider stress and arousal when investigating social attention in ASD. Future research should build on this approach by incorporating stress induction paradigms or non-invasive dual eye-tracking methodologies to explore the dynamics of social gaze and its responsiveness to pharmacological interventions in real-world settings.

Acknowledgments

We would like to thank the CoBeNe doctoral school and the University of Vienna for providing funds for this project, the CogSciHub-SPP team (David Cserjan) for assisting with recruitment, our master students who assisted in data collection and preprocessing (Ronja Lipp, Chiara Vollmerhausen, Natalie Zeilinger, Mehmet Uzunoglu, Rudolf Brozmann, Yanick Lahaye, Felix Servus, Felix Puster), Daniel Gugerell for providing the eye-tracking equipment and all participants who took part in the study.

Declaration of AI use

Generative AI (ChatGPT, model 4o) was used for language editing to improve clarity and flow, as well as for refining code. However, it was not used to generate research ideas or draw scientific conclusions.

Conflicts of Interest

The authors declare no conflicts of interest.

References

Auyeung, B., Lombardo, M. V., Heinrichs, M., Chakrabarti, B., Sule, A., Deakin, J. B., Bethlehem, R. A. I., Dickens, L., Mooney, N., Sipple, J. A. N., Thiemann, P., & Baron-Cohen, S. (2015). Oxytocin increases eye contact during a real-time, naturalistic social interaction in males with and without autism. Translational Psychiatry, 5(2). https://doi.org/10.1038/tp.2014.146

Beck, A. T., Ward, C., Mendelson, M., Mock, J., & Erbauch, J. (1978). Beck depression inventory (BDI). San Antonio, TX.

Bicknell, R. J., & Leng, G. (1982). Endogenous opiates regulate oxytocin but not vasopressin secretion from the neurohypophysis. Nature, 298(5870), 161–162.

Bouvard, M. P., Leboyer, M., Launay, J.-M., Recasens, C., Plumet, M.-H., Waller-Perotte, D., Tabuteau, F., Bondoux, D., Dugas, M., Lensing, P., & Panksepp, J. (1995). Low-dose naltrexone effects on plasma chemistries and clinical symptoms in autism: A double-blind, placebo-controlled study. Psychiatry Research, 58(3), 191–201. https://doi.org/10.1016/0165-1781(95)02601-R

Bruchas, M. R., Land, B. B., & Chavkin, C. (2010). The dynorphin/kappa opioid system as a modulator of stress-induced and pro-addictive behaviors. Brain Research, 1314, 44–55. https://doi.org/10.1016/j.brainres.2009.08.062

Buehler, R., Silani, G., & Ansorge, U. (2024). Social attention in the wild—Fixations to the eyes and autistic traits during a naturalistic interaction in a healthy sample. https://doi.org/10.21203/rs.3.rs-4016436/v1

Cañigueral, R., & Hamilton, A. F. de C. (2019). The role of eye gaze during natural social interactions in typical and autistic people. Frontiers in Psychology, 10, 560.

Cañigueral, R., Hamilton, A., & Ward, J. A. (2018). Don't Look at Me, I'm Wearing an Eyetracker! Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers, 994–998. https://doi.org/10.1145/3267305.3274123

Cardoso, C., Ellenbogen, M. A., Orlando, M. A., Bacon, S. L., & Joober, R. (2013). Intranasal oxytocin attenuates the cortisol response to physical stress: A dose–response study. Psychoneuroendocrinology, 38(3), 399–407. https://doi.org/10.1016/j.psyneuen.2012.07.013

Cardoso, C., Kingdon, D., & Ellenbogen, M. A. (2014). A meta-analytic review of the impact of intranasal oxytocin administration on cortisol concentrations during laboratory tasks: Moderation by method and mental health. Psychoneuroendocrinology, 49(1), 161–170. https://doi.org/10.1016/j.psyneuen.2014.07.014

Chelnokova, O., Laeng, B., Løseth, G., Eikemo, M., Willoch, F., & Leknes, S. (2016). The μ-opioid system promotes visual attention to faces and eyes. Social Cognitive and Affective Neuroscience, 11(12), 1902–1909. https://doi.org/10.1093/scan/nsw116

Chevallier, C., Parish-Morris, J., Mcvey, A., Rump, K. M., Sasson, N. J., Herrington, J. D., & Schultz, R. T. (2015). Measuring social attention and motivation in autism spectrum disorder using eye-tracking: Stimulus type matters. Autism Research, 8(5), 620–628. https://doi.org/10.1002/aur.1479

Chita-Tegmark, M. (2016). Social attention in ASD: A review and meta-analysis of eye-tracking studies. Research in Developmental Disabilities, 48, 79–93. https://doi.org/10.1016/j.ridd.2015.10.011

Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. British Journal of Clinical Psychology, 43(3), 245–265. https://doi.org/10.1348/0144665031752934

Cuve, H. C., Gao, Y., & Fuse, A. (2018). Is it avoidance or hypoarousal? A systematic review of emotion recognition, eye-tracking, and psychophysiological studies in young adults with autism spectrum conditions. Research in Autism Spectrum Disorders, 55, 1–13. https://doi.org/10.1016/j.rasd.2018.07.002

Dal Monte, O., Piva, M., Anderson, K. M., Tringides, M., Holmes, A. J., & Chang, S. W. C. (2017). Oxytocin under opioid antagonism leads to supralinear enhancement of social attention. Proceedings of the National Academy of Sciences of the United States of America, 114(20), 5247–5252. https://doi.org/10.1073/pnas.1702725114

Dal Monte, O., Piva, M., Morris, J. A., & Chang, S. W. C. (2016). Live interaction distinctively shapes social gaze dynamics in rhesus macaques. Journal of Neurophysiology, 116(4), 1626–1643. https://doi.org/10.1152/jn.00442.2016

Davis, M. H. (1980). Interpersonal reactivity index. https://psycnet.apa.org/doiLanding?doi=10.1037/t01093-000

Engel, S., Klusmann, H., Ditzen, B., Knaevelsrud, C., & Schumacher, S. (2019). Menstrual cycle-related fluctuations in oxytocin concentrations: A systematic review and meta-analysis. Frontiers in Neuroendocrinology, 52(August 2018), 144–155. https://doi.org/10.1016/j.yfrne.2018.11.002

Fabre-Nys, C., Meller, R. E., & Keverne, E. B. (1982). Opiate antagonists stimulate affiliative behaviour in monkeys. Pharmacology Biochemistry and Behavior, 16(4), 653–659. https://doi.org/10.1016/0091-3057(82)90432-4

Fan, S., Weinberg-Wolf, H., Piva, M., Dal Monte, O., & Chang, S. W. C. (2020). Combinatorial Oxytocin Neuropharmacology in Social Cognition. Trends in Cognitive Sciences, 24(1), 8–12.

Frazier, T. W., Strauss, M., Klingemier, E. W., Zetzer, E. E., Hardan, A. Y., Eng, C., & Youngstrom, E. A. (2017). A Meta-Analysis of Gaze Differences to Social and Nonsocial Information Between Individuals With and Without Autism. Journal of the American Academy of Child & Adolescent Psychiatry, 56(7), 546–555. https://doi.org/10.1016/j.jaac.2017.05.005

Freitag, C. M., Konrad, C., Häberlen, M., Kleser, C., von Gontard, A., Reith, W., Troje, N. F., & Krick, C. (2008). Perception of biological motion in autism spectrum disorders. Neuropsychologia, 46(5), 1480–1494. https://doi.org/10.1016/j.neuropsychologia.2007.12.025

Freitag, C. M., Retz-Junginger, P., Retz, W., Seitz, C., Palmason, H., Meyer, J., Rösler, M., & von Gontard, A. (2007). Evaluation der deutschen Version des Autismus-Spektrum-Quotienten (AQ)—Die Kurzversion AQ-k. Zeitschrift Für Klinische Psychologie Und Psychotherapie, 36(4), 280–289. https://doi.org/10.1026/1616-3443.36.4.280

Gillberg, C., Terenius, L., Hagberg, B., Witt-Engerström, I., & Eriksson, I. (1990). CSF beta-endorphins in childhood neuropsychiatric disorders. Brain and Development, 12(1), 88–92.

Gonzalez, J. P., & Brogden, R. N. (1988). Naltrexone: A Review of its Pharmacodynamic and Pharmacokinetic Properties and Therapeutic Efficacy in the Management of Opioid Dependence. Drugs, 35(3), 192–213. https://doi.org/10.2165/00003495-198835030-00002

Gossen, A., Hahn, A., Westphal, L., Prinz, S., Schultz, R. T., Gründer, G., & Spreckelmeyer, K. N. (2012). Oxytocin plasma concentrations after single intranasal oxytocin administration – A study in healthy men. Neuropeptides, 46(5), 211–215. https://doi.org/10.1016/j.npep.2012.07.001

Gregory, S. G., Connelly, J. J., Towers, A. J., Johnson, J., Biscocho, D., Markunas, C. A., Lintas, C., Abramson, R. K., Wright, H. H., Ellis, P., Langford, C. F., Worley, G., Delong, G. R., Murphy, S. K., Cuccaro, M. L., Persico, A., & Pericak-Vance, M. A. (2009). Genomic and epigenetic evidence for oxytocin receptor deficiency in autism. BMC Medicine, 7(1), Article 1. https://doi.org/10.1186/1741-7015-7-62

Guastella, A. J., Gray, K. M., Rinehart, N. J., Alvares, G. A., Tonge, B. J., Hickie, I. B., Keating, C. M., Cacciotti-Saija, C., & Einfeld, S. L. (2015). The effects of a course of intranasal oxytocin on social behaviors in youth diagnosed with autism spectrum disorders: A randomized controlled trial. Journal of Child Psychology and Psychiatry, 56(4), 444–452. https://doi.org/10.1111/jcpp.12305

Guastella, A. J., Mitchell, P. B., & Dadds, M. R. (2008). Oxytocin Increases Gaze to the Eye Region of Human Faces. Biological Psychiatry, 63(1), 3–5. https://doi.org/10.1016/j.biopsych.2007.06.026

Heimberg, R. G., Mueller, G. P., Holt, C. S., Hope, D. A., & Liebowitz, M. R. (1992). Assessment of anxiety in social interaction and being observed by others: The social interaction anxiety scale and the Social Phobia Scale. Behavior Therapy, 23(1), 53–73. https://doi.org/10.1016/S0005-7894(05)80308-9

Heinrichs, M., Baumgartner, T., Kirschbaum, C., & Ehlert, U. (2003). Social support and oxytocin interact to suppress cortisol and subjective responses to psychosocial stress. Biological Psychiatry, 54(12), 1389–1398. https://doi.org/10.1016/S0006-3223(03)00465-7

Inagaki, T. K., Ray, L. A., Irwin, M. R., Way, B. M., & Eisenberger, N. I. (2016). Opioids and social bonding: Naltrexone reduces feelings of social connection. Social Cognitive and Affective Neuroscience, 11(5), 728–735. https://doi.org/10.1093/scan/nsw006

Kalin, N. H., Shelton, S. E., & Barksdale, C. M. (1988). Opiate modulation of separation-induced distress in non-human primates. Brain Research, 440(2), 285–292. https://doi.org/10.1016/0006-8993(88)90997-3

Kassner, M., Patera, W., & Bulling, A. (2014). Pupil. Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, 1151–1160. https://doi.org/10.1145/2638728.2641695

Keverne, E. B., Martensz, N. D., & Tuite, B. (1989). Beta-endorphin concentrations in cerebrospinal fluid of monkeys are influenced by grooming relationships. Psychoneuroendocrinology, 14(1), 155–161. https://doi.org/10.1016/0306-4530(89)90065-6

King, A. (2002). Hypothalamic-Pituitary-Adrenocortical (HPA) Axis Response and Biotransformation of Oral Naltrexone Preliminary Examination of Relationship to Family History of Alcoholism. Neuropsychopharmacology, 26(6), 778–788. https://doi.org/10.1016/S0893-133X(01)00416-X

Knoll, A. T., Muschamp, J. W., Daws, S. E., Ferguson, D., Dietz, D. M., Meloni, E. G., Carroll, F. I., Nestler, E. J., Konradi, C., & Carlezon, W. A. (2011). Kappa Opioid Receptor Signaling in the Basolateral Amygdala Regulates Conditioned Fear and Anxiety in Rats. Biological Psychiatry, 70(5), 425–433. https://doi.org/10.1016/j.biopsych.2011.03.017

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. Nature, 435(7042), 673–676. https://doi.org/10.1038/nature03701

Kumsta, R., & Heinrichs, M. (2013). Oxytocin, stress and social behavior: Neurogenetics of the human oxytocin system. Current Opinion in Neurobiology, 23(1), 11–16. https://doi.org/10.1016/j.conb.2012.09.004

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. Journal of Statistical Software, 82(13). https://doi.org/10.18637/jss.v082.i13

Lalanne, L., Ayranci, G., Filliol, D., Gavériaux-Ruff, C., Befort, K., Kieffer, B. L., & Lutz, P.-E. (2017). Kappa opioid receptor antagonism and chronic antidepressant treatment have beneficial activities on social interactions and grooming deficits during heroin abstinence. Addiction Biology, 22(4), 1010–1021. https://doi.org/10.1111/adb.12392

Leboyer, M., Bouvard, M. P., Launay, J.-M., Tabuteau, F., Waller, D., Dugas, M., Kerdelhue, B., Lensing, P., & Panksepp, J. (1992). Brief report: A double-blind study of naltrexone in infantile autism. Journal of Autism and Developmental Disorders, 22(2), 309–319. https://doi.org/10.1007/BF01058158

Leboyer, M., Philippe, A., Bouvard, M., Guilloud-Bataille, M., Bondoux, D., Tabuteau, F., Feingold, J., Mouren-Simeoni, M.-C., & Launay, J.-M. (1999). Whole blood serotonin and plasma beta-endorphin in autistic probands and their first-degree relatives. Biological Psychiatry, 45(2), 158–163. https://doi.org/10.1016/S0006-3223(97)00532-5

Lehrl, S. (1999). Mehrfachwahl-Wortschatz-Intelligenztest: MWT-B. Spitta.

Lenth, R. (2022). emmeans: Estimated marginal means, aka least-squares means. R package version 1.7. 2.

LoParo, D., & Waldman, I. D. (2015). The oxytocin receptor gene (OXTR) is associated with autism spectrum disorder: A meta-analysis. Molecular Psychiatry, 20(5), 640–646.

Løseth, G. E., Ellingsen, D.-M., & Leknes, S. (2014). State-dependent μ-opioid modulation of social motivation. Frontiers in Behavioral Neuroscience, 8, 430.

Løseth, G., Trøstheim, M., & Leknes, S. (2024). Endogenous mu-opioid modulation of social connection in humans: A systematic review and meta-analysis. Translational Psychiatry, 14(1), 1–11. https://doi.org/10.1038/s41398-024-03088-3

Lukas, M., Toth, I., Reber, S. O., Slattery, D. A., Veenema, A. H., & Neumann, I. D. (2011). The Neuropeptide Oxytocin Facilitates Pro-Social Behavior and Prevents Social Avoidance in Rats and Mice. Neuropsychopharmacology, 36(11), 2159–2168. https://doi.org/10.1038/npp.2011.95

Macdonald, R. G., & Tatler, B. W. (2018). Gaze in a real-world social interaction: A dual eye-tracking study. Quarterly Journal of Experimental Psychology, 71(10), 2162–2173. https://doi.org/10.1177/1747021817739221

Machin, A. J., & Dunbar, R. I. M. (2011). The brain opioid theory of social attachment: A review of the evidence. https://doi.org/10.1163/000579511X596624

Martel, F. L., Nevison, C. M., Rayment, F. D., Simpson, M. J. A., & Keverne, E. B. (1993). Opioid receptor blockade reduces maternal affect and social grooming in rhesus monkeys. Psychoneuroendocrinology, 18(4), 307–321. https://doi.org/10.1016/0306-4530(93)90027-I

Martins, D., Brodmann, K., Veronese, M., Dipasquale, O., Mazibuko, N., Schuschnig, U., Zelaya, F., Fotopoulou, A., & Paloyelis, Y. (2022). "Less is more": A dose-response account of intranasal oxytocin pharmacodynamics in the human brain. Progress in Neurobiology, September 2021, 102239. https://doi.org/10.1016/j.pneurobio.2022.102239

Mayrand, F., Capozzi, F., & Ristic, J. (2023). A dual mobile eye tracking study on natural eye contact during live interactions. Scientific Reports, 13(1), 11385.

McLaughlin, J. P., Li, S., Valdez, J., Chavkin, T. A., & Chavkin, C. (2006). Social Defeat Stress-Induced Behavioral Responses are Mediated by the Endogenous Kappa Opioid System. Neuropsychopharmacology, 31(6), 1241–1248. https://doi.org/10.1038/sj.npp.1300872

Miller, R., Wojtyniak, J.-G., Weckesser, L. J., Alexander, N. C., Engert, V., & Lehr, T. (2018). How to disentangle psychobiological stress reactivity and recovery: A comparison of model-based and non-compartmental analyses of cortisol concentrations.

Psychoneuroendocrinology, 90, 194–210. https://doi.org/10.1016/j.psyneuen.2017.12.019

Milton, D. E. M. (2012). On the ontological status of autism: The 'double empathy problem.' Disability & Society, 27(6), 883–887. https://doi.org/10.1080/09687599.2012.710008

Modahl, C., Green, L. A., Fein, D., Morris, M., Waterhouse, L., Feinstein, C., & Levin, H. (1998). Plasma oxytocin levels in autistic children. Biological Psychiatry, 43(4), 270–277. https://doi.org/10.1016/S0006-3223(97)00439-3

Moriuchi, J. M., Klin, A., & Jones, W. (2017). Mechanisms of Diminished Attention to Eyes in Autism. American Journal of Psychiatry, 174(1), 26–35. https://doi.org/10.1176/appi.ajp.2016.15091222

Nagamitsu, S., Matsuishi, T., Kisa, T., Komori, H., Miyazaki, M., Hashimoto, T., Yamashita, Y., Ohtaki, E., & Kato, H. (1997). CSF β-Endorphin Levels in Patients with Infantile Autism. Journal of Autism and Developmental Disorders, 27(2), 155–163. https://doi.org/10.1023/A:1025839807431

Neumann, I. D., & Landgraf, R. (2012). Balance of brain oxytocin and vasopressin: Implications for anxiety, depression, and social behaviors. Trends in Neurosciences, 35(11), 649–659. https://doi.org/10.1016/j.tins.2012.08.004

Nummenmaa, L., Tuominen, L., Dunbar, R., Hirvonen, J., Manninen, S., Arponen, E., Machin, A., Hari, R., Jääskeläinen, I. P., & Sams, M. (2016). Social touch modulates endogenous μ-opioid system activity in humans. NeuroImage, 138, 242–247. https://doi.org/10.1016/j.neuroimage.2016.05.063

Ooi, Y. P., Weng, S.-J., Kossowsky, J., Gerger, H., & Sung, M. (2016). Oxytocin and Autism Spectrum Disorders: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. Pharmacopsychiatry, 50, 5–13. https://doi.org/10.1055/s-0042-109400

Panksepp, J. (1979). A neurochemical theory of autism. Trends in Neurosciences, 2, 174–177.

Pellissier, L. P., Gandía, J., Laboute, T., Becker, J. A. J., & Le Merrer, J. (2018). μ opioid receptor, social behaviour and autism spectrum disorder: Reward matters. British Journal of Pharmacology, 175(14), 2750–2769. https://doi.org/10.1111/bph.13808

Pruessner, J. C., Kirschbaum, C., Meinlschmid, G., & Hellhammer, D. H. (2003). Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. Psychoneuroendocrinology, 28(7), 916–931. https://doi.org/10.1016/S0306-4530(02)00108-7

Putnam, P. T., & Chang, S. W. (2022a). Interplay between the oxytocin and opioid systems in regulating social behaviour. Philosophical Transactions of the Royal Society B, 377(1858), 20210050.

Putnam, P. T., & Chang, S. W. C. (2022b). Oxytocin does not stand alone. Philosophical Transactions of the Royal Society B: Biological Sciences, 377(1858), 20210047. https://doi.org/10.1098/rstb.2021.0047

Quintana, D. S. (2018). Revisiting non-significant effects of intranasal oxytocin using equivalence testing. Psychoneuroendocrinology, 87, 127–130. https://doi.org/10.1016/j.psyneuen.2017.10.010

Quintana, D. S., Lischke, A., Grace, S., Scheele, D., Ma, Y., & Becker, B. (2021). Advances in the field of intranasal oxytocin research: Lessons learned and future directions for clinical research. Molecular Psychiatry, 26(1), 80–91. https://doi.org/10.1038/s41380-020-00864-7

Quintana, D. S., Westlye, L. T., Alnæs, D., Rustan, Ø. G., Kaufmann, T., Smerud, K. T., Mahmoud, R. A., Djupesland, P. G., & Andreassen, O. A. (2016). Low dose intranasal oxytocin delivered with Breath Powered device dampens amygdala response to emotional stimuli: A peripheral effect-controlled within-subjects randomized dose-response fMRI trial. Psychoneuroendocrinology, 69, 180–188. https://doi.org/10.1016/j.psyneuen.2016.04.010

Quintana, D. S., Westlye, L. T., Rustan, Ø. G., Tesli, N., Poppy, C. L., Smevik, H., Tesli, M., Røine, M., Mahmoud, R. A., Smerud, K. T., Djupesland, P. G., & Andreassen, O. A. (2015). Low-dose oxytocin delivered intranasally with Breath Powered device affects social-cognitive behavior: A randomized four-way crossover trial with nasal cavity dimension assessment. Translational Psychiatry, 5(7), e602–e602. https://doi.org/10.1038/tp.2015.93

R Core Team. (2022). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing). https://www.R-project.org/

Rilling, J. K., DeMarco, A. C., Hackett, P. D., Thompson, R., Ditzen, B., Patel, R., & Pagnoni, G. (2012). Effects of intranasal oxytocin and vasopressin on cooperative behavior and associated brain activity in men. Psychoneuroendocrinology, 37(4), 447–461. https://doi.org/10.1016/j.psyneuen.2011.07.013

Rogers, S. L., Speelman, C. P., Guidetti, O., & Longmuir, M. (2018). Using dual eye tracking to uncover personal gaze patterns during social interaction. Scientific Reports, 8(1), 4271.

Ross, D. L., Klykylo, W. M., & Hitzemann, R. (1987). Reduction of elevated CSF beta-endorphin by fenfluramine in infantile autism. Pediatric Neurology, 3(2), 83–86. https://doi.org/10.1016/0887-8994(87)90032-4

Roy, A., Roy, M., Deb, S., Unwin, G., & Roy, A. (2015). Are opioid antagonists effective in attenuating the core symptoms of autism spectrum conditions in children: A systematic review. Journal of Intellectual Disability Research, 59(4), 293–306. https://doi.org/10.1111/jir.12122

Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., & Baron-Cohen, S. (2015). Measuring autistic traits in the general population: A systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. Molecular Autism, 6, 2. https://doi.org/10.1186/2040-2392-6-2

Sandi, C., & Haller, J. (2015). Stress and the social brain: Behavioural effects and neurobiological mechanisms. Nature Reviews Neuroscience, 16(5), 290–304. https://doi.org/10.1038/nrn3918

Senju, A., & Johnson, M. H. (2009). Atypical eye contact in autism: Models, mechanisms and development. Neuroscience & Biobehavioral Reviews, 33(8), 1204–1214. https://doi.org/10.1016/j.neubiorev.2009.06.001

Shamay-Tsoory, S. G., & Abu-Akel, A. (2016). The social salience hypothesis of oxytocin. Biological Psychiatry, 79(3), 194–202.

Spielberger, C. D., Gonzalez-Reigosa, F., Martinez-Urrutia, A., Natalicio, L. F., & Natalicio, D. S. (1971). The state-trait anxiety inventory. Revista Interamericana de Psicologia/Interamerican Journal of Psychology, 5(3 & 4). http://journal.sipsych.org/index.php/IJP/article/view/620

Stuart, N., Whitehouse, A., Palermo, R., Bothe, E., & Badcock, N. (2022). Eye gaze in autism spectrum disorder: A review of neural evidence for the eye avoidance hypothesis. Journal of Autism and Developmental Disorders, 1–22.

Taylor, J. L., & Corbett, B. A. (2014). A review of rhythm and responsiveness of cortisol in individuals with autism spectrum disorders. Psychoneuroendocrinology, 49, 207–228. https://doi.org/10.1016/j.psyneuen.2014.07.015

Tordjman, S., Anderson, G. M., Botbol, M., Brailly-Tabard, S., Perez-Diaz, F., Graignic, R., Carlier, M., Schmit, G., Rolland, A.-C., Bonnot, O., Trabado, S., Roubertoux, P., & Bronsard, G. (2009). Pain Reactivity and Plasma β-Endorphin in Children and Adolescents with

Autistic Disorder. PLOS ONE, 4(8), e5289. https://doi.org/10.1371/journal.pone.0005289

Troisi, A., Frazzetto, G., Carola, V., Di Lorenzo, G., Coviello, M., D'Amato, F. R., Moles, A., Siracusano, A., & Gross, C. (2011). Social hedonic capacity is associated with the A118G polymorphism of the mu-opioid receptor gene (OPRM1) in adult healthy volunteers and psychiatric patients. Social Neuroscience, 6(1), 88–97. https://doi.org/10.1080/17470919.2010.482786

Vandesande, F., & Dierickx, K. (1975). Identification of the vasopressin producing and of the oxytocin producing neurons in the hypothalamic magnocellular neurosecretory system of the rat. Cell and Tissue Research, 164(2), 153–162.

Vatheuer, C. C., Vehlen, A., Von Dawans, B., & Domes, G. (2021). Gaze behavior is associated with the cortisol response to acute psychosocial stress in the virtual TSST. Journal of Neural Transmission, 128(9), 1269–1278. https://doi.org/10.1007/s00702-021-02344-w

Vehlen, A., Spenthof, I., Tönsing, D., Heinrichs, M., & Domes, G. (2021). Evaluation of an eye tracking setup for studying visual attention in face-to-face conversations. Scientific Reports, 11(1), 2661. https://doi.org/10.1038/s41598-021-81987-x

Vuong, C., Van Uum, S. H. M., O'Dell, L. E., Lutfy, K., & Friedman, T. C. (2010). The effects of opioids and opioid analogs on animal and human endocrine systems. Endocrine Reviews, 31(1), 98–132. https://doi.org/10.1210/er.2009-0009

Wardle, M. C., Bershad, A. K., & de Wit, H. (2016). Naltrexone alters the processing of social and emotional stimuli in healthy adults. Social Neuroscience, 11(6), 579–591. https://doi.org/10.1080/17470919.2015.1136355

Way, B. M., Taylor, S. E., & Eisenberger, N. I. (2009). Variation in the μ-opioid receptor gene (OPRM1) is associated with dispositional and neural sensitivity to social rejection. Proceedings of the National Academy of Sciences, 106(35), 15079–15084. https://doi.org/10.1073/pnas.0812612106

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. https://ggplot2.tidyverse.org

Windle, R. J., Shanks, N., Lightman, S. L., & Ingram, C. D. (1997). Central Oxytocin Administration Reduces Stress-Induced Corticosterone Release and Anxiety Behavior in Rats*. Endocrinology, 138(7), 2829–2834. https://doi.org/10.1210/endo.138.7.5255

Young, L. J., & Barrett, C. E. (2015). Can oxytocin treat autism? Science, 347(6224), 825–826. https://doi.org/10.1126/science.aaa8120

Young, L. J., & Wang, Z. (2004). The neurobiology of pair bonding. Nature Neuroscience, 7(10), 1048–1054. https://doi.org/10.1038/nn1327

6 Supplementary Material

Fixation Duration

As an alternative approach to our preprocessing and analysis of eye tracking data, we also investigated an alternative dependent variable, specifically the percentage of total fixation duration on each AOI. This analysis mirrored the main analysis reported in the main paper: A significant group by AOI interaction was observed, $F(3,979) = 8.35$, $p < 0.001$, with post-hoc tests indicating significant group differences for the eye ($t(619) = -2.86$, $p = 0.004$) and background regions ($t(618) = 3.92$, $p < 0.001$). No interaction with medication was observed ($p > 0.05$), but group differences were again most prominent for the PLA/PLA condition, where ASD participants fixated significantly less on the eyes compared to controls ($t(979) = -2.30$, p = 0.022).

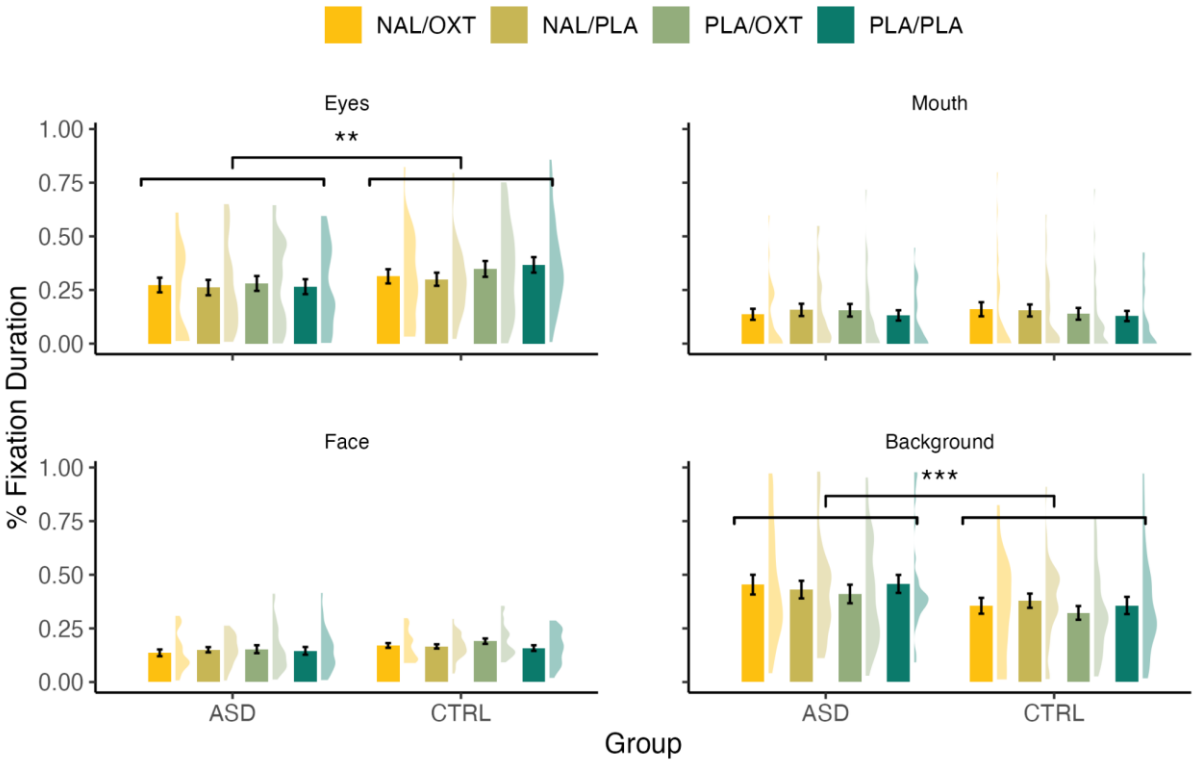

Figure S1 - Fixation rates on different areas of interest (AOIs) across menstrual phases in the ASD and control (CTRL) groups. AOIs include the eyes, mouth, face, and background. Fixation patterns varied slightly across phases, but no significant main effects or interactions with menstrual phase were observed. Error bars represent standard errors of the mean (SEM).

Menstrual Phase (SM1)

Running an LMM with menstrual phase as an additional predictor (model S1) did not reveal any significant main ($F(3, 246) = 0.248$, $p = 0.86$) or interactive effects (all $p > 0.05$) of menstrual phase with group or medication. Similar to the main analyses, the only significant effects were AOI ($F(3, 232) = 53.4$, $p < 0.0001$) and the AOI × Group interaction ($F(3, 232) = 2.927$, $p = 0.035$). This suggests that different phases of the menstrual cycle did not significantly influence fixations on the eyes, nor did they impact medication effects. In contrast, group differences in fixations on the eyes and background when collapsing over menstrual phases were significant (see Figure S1).

In summary, the control analysis with menstrual cycle suggested no effect of menstrual cycle phases on group differences or medication effects. Group differences for the eye region were significant, similar to our main analysis. Thus, although fluctuations of oxytocin over the menstrual cycle have been reported, they did not significantly increase or decrease the effect of oxytocin in our study.

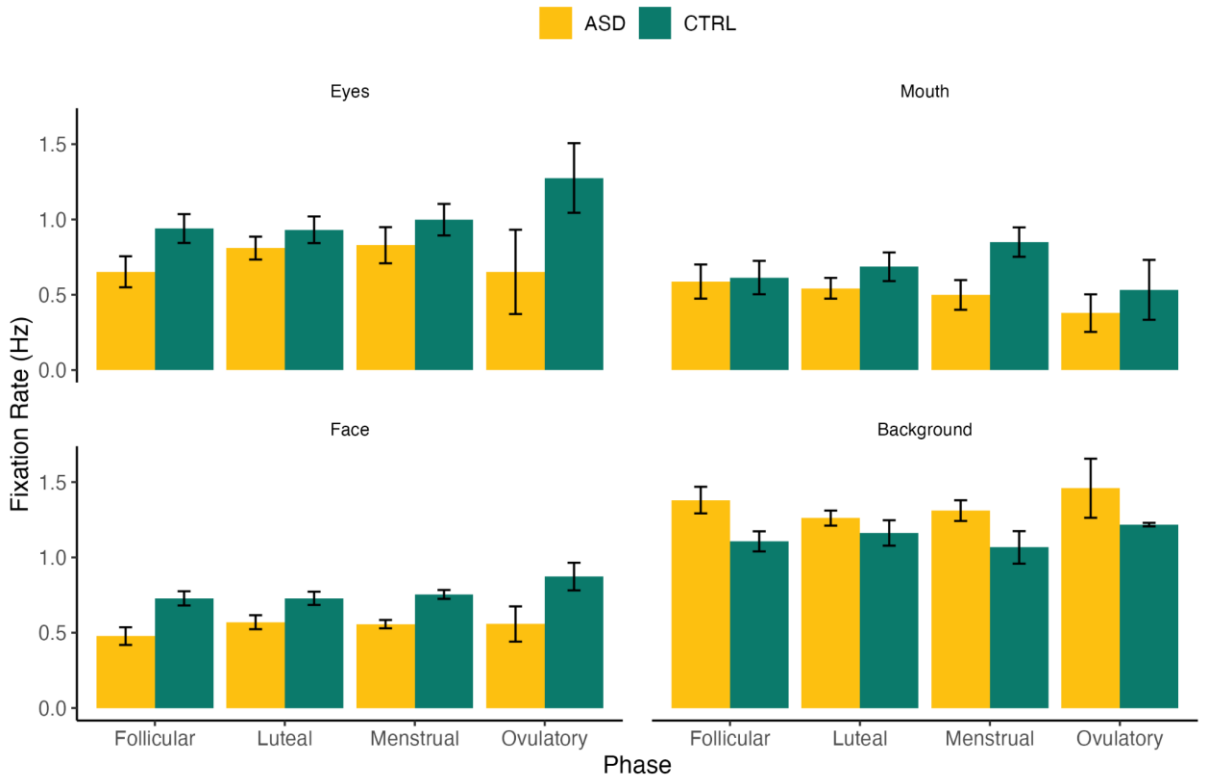

Figure S2 - Fixation rates on different areas of interest (AOIs) across menstrual phases in the ASD and control (CTRL) groups. AOIs include the eyes, mouth, face, and background. Fixation patterns varied slightly across phases, but no significant main effects or interactions with menstrual phase were observed. Error bars represent standard errors of the mean (SEM).

Sympathy ratings

Sympathy ratings were measured using a questionnaire where participants evaluated the experimenter on four dimensions. One item was used for each dimension. Specifically, participants were asked whether the experimenter seemed to be an interesting, good, kind and open-minded person. Each item was rated on a 7-point scale, ranging from "not at all" to "very much." For use as a control variable in statistical analyses, the mean of these measures over all sessions was used as a measure of sympathy, i.e. how sympathetic the experimenter was perceived during the experiment.

Fixation Time Course

For exploratory purposes, longitudinal plots of fixations were created to investigate the time course of fixations to AOIs for each subject in real-time. These plots suggest relative stability of gaze behavior over interview time for a majority of subjects. An exemplary plot for a single ASD subject is provided in Figure S3.

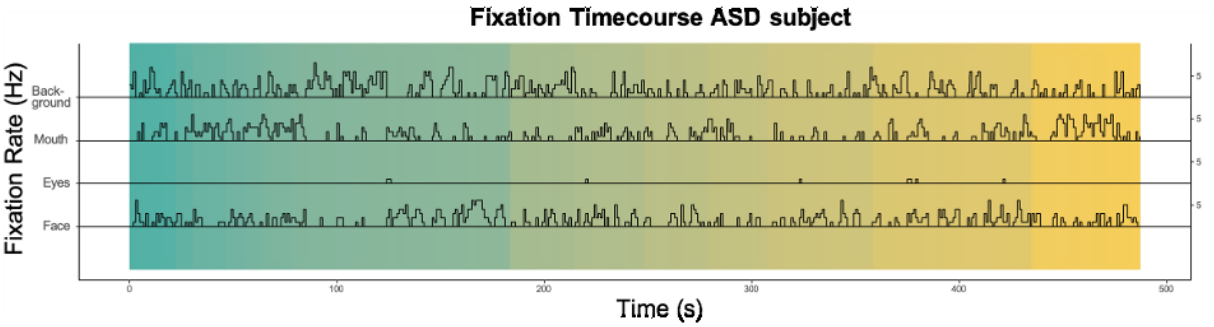

Figure S3 - Longitudinal plot of a single ASD subject, depicting fixations rates over time on different AOIs. The subject is rarely fixating on the eyes during the entire interview. Fixation rates were binned at 1s time intervals.

Interview Questions

Instructions for the participant (guideline):

"We are now going to have a short conversation. Please make yourself comfortable, so you will be able to sit in the same position for the next few minutes. Please try to avoid large movements of your head, but small movements like nodding are ok. I will ask you a couple of questions and later on, you will have the opportunity to ask me some questions as well. The idea is to have as relaxed and natural a conversation as possible. If you are ready, I will now start with the first question."

Interview Questions:

1. Did you sleep well last night?
2. Did you dream about anything?
3. Have you eaten at a restaurant in the past week?
4. Were you out in nature last week?
5. Have you read a book in the past week?
6. Did you watch a movie or TV show recently?
7. What did you do last weekend?
8. Could you tell me what you plan to do next weekend?
9. What was the most interesting event you experienced in the past week?
10. When you woke up this morning, were you looking forward to participating in this study, or did it feel more like an obligation?
11. Could you tell me about your journey to the lab, did anything interesting happen on the way?
12. How did you feel after receiving the medication?
13. Do you have any questions you would like to ask me?

General Discussion

1 Summary of Thesis Goal and Theoretical Framework

The overarching goal of this thesis was to investigate how specific neurobiological systems shape social behavior, with a particular focus on the interactive effects of opioids and oxytocin. Existing theories have either proposed that social behavior is governed by domain-specific systems, e.g. as suggested by the *social motivation* and *social brain hypotheses* (Chevallier et al., 2015; Frith, 2007) or by more general reward-related mechanisms, potentially involving the opioid system (Pellissier et al., 2018). Opioids and oxytocin interact and their effects on social behavior may partially be modulated by the hypothalamic-pituitary-adrenal (HPA) axis (Fan et al., 2020). Importantly, different opioid receptor subtypes, including μ- and κ-receptors, are implicated in both social reward and stress responses (Bruchas et al., 2010; Knoll et al., 2011), suggesting that opioid antagonism may affect social gaze not only via hedonic valuation but also through arousal-related mechanisms (Sandi & Haller, 2015). Furthermore, predictive coding theories propose that ASD involves alterations in how sensory information and prior expectations are integrated, with competing accounts emphasizing different parameters and levels of the learning hierarchy (Lawson et al., 2017). This thesis aimed to disentangle the contributions of these systems and their interactions in shaping social attention and social learning.

Another important goal of this thesis was to introduce methodological innovations across three levels. First, while reduced attention to the eye region in ASD has consistently been demonstrated in screen-based settings, experiments in naturalistic contexts remain scarce. Furthermore, learning tasks often rely on stimuli that differ substantially in low-level visual features at which differences in perceptual processing are often assumed to arise in ASD. Third, interactive pharmacological studies targeting social behavior in humans are rare. To address these gaps, this thesis comprises three empirical studies: (1) a naturalistic eye tracking study capturing gaze behavior during a realistic social interaction, combined with computer vision-based face detection algorithms (2) a computational modeling study applying a Rescorla-Wagner reinforcement learning model to choice behavior in social and non-social contexts using facial point-light displays (PLDs); and (3) a double-blind, placebo-controlled, pharmacological study involving the administration of oxytocin and naltrexone, assessing social gaze in both neurotypical individuals and participants with ASD.

By combining ecologically valid behavioral paradigms, computational modeling, and pharmacological interventions, this thesis aimed to advance our understanding of the neurobiological and computational mechanisms that govern social attention and learning, and how these mechanisms may be altered in individuals with ASD.

2 Main Findings

Chapter 1 consisted of a proof-of-concept for studying social attention in naturalistic settings using mobile eye tracking and automated face detection. In contrast to traditional screen-based paradigms, this study captured gaze behavior during a live, face-to-face interview, employing a head-mounted eye-tracking system in combination with a convolutional neural network (MTCNN) for real-time face and facial landmark detection. Participants completed a semi-structured interview while their gaze was recorded and analyzed frame-by-frame with respect to automatically defined areas of interest (AOIs), including the eyes, mouth, and background.

Analyses revealed a significant negative association between autistic traits as measured by the Autism Spectrum Quotient (AQ) and fixation rates to the eye region. This effect persisted across the duration of the

interaction and was not limited to specific question types or interaction phases. In contrast to some previous findings, no significant increase in gaze on mouth was observed for participants with high AQ scores.

By demonstrating that naturalistic gaze patterns can be quantified with minimal manual coding, this study demonstrates the feasibility and validity of combining mobile eye tracking with computer vision approaches. It provides a scalable paradigm for investigating social attention under conditions that more closely approximate real-life social interactions.

Chapter 2 examined whether autistic traits influence learning from social versus non-social feedback using a computational modeling approach. Participants completed a probabilistic reinforcement learning task in which feedback was provided through point-light displays (PLDs) depicting either facial expressions (social condition) or abstract symbols (non-social condition) that were matched on low-level visual features.

Behavioral data were modeled using a Rescorla-Wagner reinforcement learning framework to estimate latent parameters including learning rate (α) and reward sensitivity (ρ). Participants showed robust learning across both conditions. Raw performance was marginally lower in the social condition across the full sample, but did not significantly differ as a function of autistic traits. This suggests that learning from facial PLDs may have been slightly harder overall, potentially due to the challenge of recognizing emotional expressions from PLDs. Computational modeling further revealed that higher autistic traits were associated with reduced reward sensitivity (ρ) in the win domain, indicating a diminished valuation of rewarding outcomes. This effect did not significantly differ between social and non-social conditions.

As such, the findings do not provide strong evidence for the social motivation hypothesis, which would predict a trait-related reduction in reward sensitivity specifically for social stimuli. Instead, the observed attenuation in reward sensitivity may reflect a more general processing difference, possibly characterized by slower integration of reward feedback as suggested by the slow updating hypothesis. However, it is important to note that individuals with higher autistic traits may have relied on compensatory strategies employing explicit cognitive effort rather than automatic processing of emotional facial expressions. Overall, results underscore the value of computational modeling in detecting subtle individual differences and establish PLDs as a tool for disentangling domain-specific from domain-general effects in social learning.

Chapter 3 investigated the interactive effects of oxytocin and opioid antagonism on social attention employing the paradigm developed in chapter 1. Participants completed four testing sessions involving combinations of intranasal oxytocin and oral naltrexone. Salivary cortisol served as a marker of HPA axis activity.

Contrary to our hypotheses, no significant main effects of medication were observed on gaze to the eye region across the full sample. However, exploratory analyses revealed that group differences between ASD and neurotypical participants were most pronounced under placebo and appeared attenuated under pharmacological manipulation. Critically, cortisol responses to naltrexone were differentially related to social gaze between groups: in the ASD group, higher cortisol levels were associated with increased fixations to the eyes, whereas in controls, the opposite pattern was observed. These opposing associations point to different stress- or arousal- related mechanisms under opioid antagonism in ASD versus neurotypical individuals.

This study represents the first attempt to simultaneously target the opioid and oxytocin systems in humans. While behavioral gaze outcomes remained largely stable across medication conditions, a robust group difference in social gaze emerged. This effect mirrors findings from Chapter 1, where autistic traits related to reduced fixations on eyes, and provides encouraging evidence for the sensitivity of the naturalistic paradigm developed in

this thesis. Together, these findings underscore the relevance of opioid-oxytocin-HPA axis interactions for future models of social attention in ASD and support the utility of naturalistic experimental designs in social neuroscience.

Integration Across Studies

Taken together, the three studies in this thesis converge on a consistent pattern: individual differences in autistic traits and clinical diagnoses of ASD are reliably associated with reduced eye-directed gaze during naturalistic social interaction, which validates the eye-tracking paradigm developed here as a sensitive tool for capturing social gaze. In contrast, Chapter 2 revealed no selective impairments in social learning performance, but computational modeling uncovered a subtle reduction in reward sensitivity among individuals with higher autistic traits. This finding can be interpreted as suggestive of domain-general processes, rather than selective social motivational differences underlying autistic traits. Consequently, reduced eye gaze in ASD may not be fully attributable to diminished social reward value but could instead arise from broader alterations in reward sensitivity. However, further research explicitly testing this assumption is necessary before making final conclusions.

The interpretation that reduced social gaze in ASD does not primarily reflect a deficit in social motivation is further supported by findings from Chapter 3. Although pharmacological manipulation of the oxytocin and opioid systems did not yield robust main effects on gaze behavior, it elicited distinct cortisol responses that interacted with social gaze differently across groups. These results highlight the role of stress- and arousal-related mechanisms in shaping social attention and suggest that, in naturalistic settings, social gaze may not be exclusively governed by reward-related processes. Notably, although large group differences in social anxiety were observed, these did not mediate social gaze. This can be regarded as an argument against the notion of hyperarousal-induced gaze avoidance in ASD. Instead, the results are more consistent with a model of gaze indifference, modulated not only by motivational aspects, but also by optimal levels of arousal. This perspective aligns with theories emphasizing the initial state of the individual and the interplay of reward-processing with other systems, proposed by the state-dependent SOMSOM model by Løseth et al and by the μ-receptor balance model by Pellissier and colleagues.

3 Theoretical Implications and Scientific Contributions

Social Motivation Hypothesis

The present findings provide only limited support for the social motivation hypothesis. While reduced eye-directed gaze was consistently observed in relation to autistic traits and diagnostic status, the underlying mechanisms identified here suggest that this reduction may not stem from a selective deficit in social reward processing. In Chapter 2, reduced reward sensitivity associated with autistic traits did not interact with the social nature of the feedback stimuli, arguing against a domain-specific motivational deficit. Similarly, in Chapter 3, pharmacological manipulation of the oxytocin and opioid systems did not elicit main effects on social gaze behavior, further weakening the notion that neurobiological systems associated with social reward are the primary driver of atypical gaze patterns in ASD. Instead, results point to a broader, domain-general alteration in reward sensitivity and arousal regulation mechanisms. Although we do not refute the relevance of social reward processing, our results suggest that models focusing exclusively on motivation may overlook important modulatory processes, particularly in naturalistic contexts, where stress and arousal-related mechanisms may exert an important influence.

Predictive Coding Framework

The results also have implications for predictive coding accounts of ASD, which propose that autistic individuals may process sensory information and priors differently, leading to altered learning and adaptation. While the current experiments were not designed to model hierarchical precision weighting, the behavioral and computational data from Chapter 2 are more consistent with the slow updating hypothesis than with the notion of weak priors. Specifically, the reduced reward sensitivity observed in individuals with higher autistic traits may reflect less flexible integration of feedback over time, in line with accounts positing slower error correction and diminished trial-by-trial updating. Notably, this effect emerged across both social and non-social learning conditions, lending support to domain-general, rather than domain-specific, predictive processing differences in ASD.

Opioid and Oxytocin Interactions

A central aim of this thesis was to investigate how oxytocin and opioids interact to shape social behavior. While previous studies have often focused on these systems in isolation, the current work represents one of the first attempts to assess their combined effects in humans. The findings from Chapter 3 did not reveal direct effects of either oxytocin or opioid antagonism on gaze behavior. However, exploratory analyses uncovered cortisol responses under naltrexone and opposing associations with cortisol across groups. These differential patterns suggest that the opioidergic system may partially influence social gaze via its role in stress- and arousal regulation. Importantly, baseline social gaze differences (mirroring those in Chapter 1) were attenuated by pharmacological modulation. This attenuation may reflect a shift in internal state, suggesting that individual differences in stress or arousal levels play a role in shaping the behavioral impact of pharmacological interventions. This interpretation aligns with theoretical frameworks that emphasize *state-dependency*, such as the SOMSOM model and the μ-balance model, which propose that the effects of opioid system activity on social behavior are dependent upon the individual's initial physiological or motivational state. For example, the SOMSOM model predicts that when in distress, opioid antagonists exacerbate distress and increase social seeking, while the same manipulation leads to

a state of relative indifference and reduced social exploration when in a neutral state. The balance model by Pellissier et al. further suggests that both excessive and deficient opioid activity can disrupt the balance between approach and avoidance systems, particularly under stress, where a hyperactive avoidance system may dominate and suppress social exploration. Importantly, our results may suggest different baseline states across groups during the experiment: In the ASD group, higher cortisol reactivity was associated with increased fixations on the eyes, possibly indicating a shift from a hypoaroused or socially indifferent baseline to a more optimal level of arousal that facilitated social attention. In contrast, in the control group, higher cortisol was linked to reduced eye contact, consistent with stress-induced gaze avoidance and a potential state of hyperarousal. While this interpretation remains speculative, the results presented here point to differential social effects of opioid antagonism across groups via stress and avoidance, which should be further explored in future studies.

Methodological Innovations

This thesis also introduces several methodological advances. First, a novel naturalistic eye-tracking paradigm was developed, involving semi-structured face-to-face interviews paired with computer vision-based face detection. This approach enabled quantification of gaze behavior to socially relevant areas of interest (AOIs) without the need for labor-intensive manual coding. This approach advances the methodology in social neuroscience towards more naturalistic experimental designs, offering a scalable method to capture gaze behavior in real-life social contexts, an essential step for identifying social differences in ASD that may not emerge under artificial laboratory conditions. Second, the reinforcement learning task employed point-light displays (PLDs), allowing for the comparison of social and non-social conditions while matching stimuli on low-level visual features. This controlled for perceptual confounds often present in previous studies and enabled a more precise assessment of domain-specific versus domain-general learning processes. Third, the pharmacological study combined the administration of oxytocin and the opioid antagonist naltrexone in a double-blind, placebo-controlled design. To our knowledge, this represents the first attempt to simultaneously explore oxytocinergic and opioidergic mechanisms.

4 Limitations and Future Directions

Eye Tracking

While the use of head-mounted eye tracking during live interaction increases ecological validity, it remains somewhat invasive and may itself alter natural gaze behavior (Cañigueral et al., 2021). However, non-invasive and reliable alternatives to head-worn eye trackers are currently limited. While promising workarounds exist for table-based eye trackers to enable face-to-face interactions and dual eye-tracking (Tönsing et al., 2025), such setups are relatively rare due to technological challenges. For example, remote eye trackers require participants to maintain a stable head position within a limited tracking area and precision degrades significantly when participants move away from optimal positioning. Consequently, most studies investigating real-life social interactions prefer head-mounted eye trackers, which offer greater flexibility and accuracy in dynamic settings. Nonetheless, further development of non-invasive, table-based solutions could prove valuable for expanding the ecological validity of social neuroscience paradigms.

Moreover, although the present paradigm captured gaze to facial regions using automated AOI tracking, it did not account for finer-grained dynamics of social interaction, such as the content of the conversation. Advances in natural language processing now also allow for content analysis, i.e. via large language models, which could be used to assess the semantic and affective content of dialogue in real time, segment the interview accordingly, and investigate gaze behavior during different topics or affective content. Similarly, emotional expression analysis via emotion recognition algorithms represents a feasible extension of the face detection algorithm presented here. Capturing the onset of emotional expressions, such as smiles, during the interaction and applying lag-sequential or temporal synchrony analyses may yield a more fine-grained understanding of the reciprocal dynamics of social interactions. Additionally, the questions used here were not drawn from well-established protocols. Future studies could benefit from adopting validated semi-structured interaction procedures, analogous to how the Trier Social Stress Test (TSST) standardizes social stress induction, with the “fast friends” protocol offering a promising framework (Tönsing et al., 2025).

Furthermore, complementing cortisol sampling with psychophysiological measures (e.g., heart rate variability, electrodermal activity) would allow to infer the participants’ autonomic state during the interaction and test how internal arousal relates to gaze behavior. Subjective reports on interaction quality may also help to address the debate on gaze aversion and indifference (Moriuchi). Prior qualitative work suggests that individuals with ASD perceive eye contact as stressful or aversive, though others may view it as neutral or uninformative (Trevisan et al., 2017). This heterogeneity highlights the importance of distinguishing potential subgroups within the autism spectrum. Finally, this study examined only unidirectional gaze behavior and did not account for the bidirectional nature of real-world social interactions. Akin to the “double empathy problem” (Milton, 2012), which underscores the importance of the alignment or mismatch between differing personality traits and cognitive styles, it is critical to consider the characteristics and behaviors of both interaction partners. Future studies may benefit from incorporating dual eye tracking and collecting data from both interaction partners, thereby enabling a truly interactive account of social attention (Schilbach et al., 2013).

Learning

The reinforcement learning task with social and non-social point-light displays (PLDs) is subject to several limitations as well. First, we did not incorporate a hierarchical Bayesian learning framework (e.g., Hierarchical Gaussian Filter) or manipulate volatility (shifts in underlying reward probability) which could have provided a more detailed account of belief updating under uncertainty. Though some hypotheses derived from the predictive processing account presume learning atypicalities in ASD at lower levels of the learning hierarchy (i.e. weak priors hypothesis, slow updating hypothesis), several studies emphasize that learning differences may occur at higher levels, for example an overestimation of the volatility of the environment (Lawson et al., 2017).

Second, while PLDs offer control over low-level visual features, they lack ecological validity and may not evoke sufficiently strong or affectively rich social cues to elicit robust group differences. This limitation may prompt participants to engage more cognitive, top-down learning strategies, potentially obscuring the influence of autistic traits on social learning. It further raises the question of whether more realistic or emotionally engaging stimuli, such as videos of facial expressions, could yield stronger effects. In the pharmacological experiment presented here, data was also collected using a learning task featuring realistic facial expressions and complex, fractal-like stimuli as social and non-social feedback. Analysis of this data is scheduled for the near future and may help address this question.

Third, linking computational learning parameters (e.g., learning rates, reward sensitivity) to real-life social attention remains difficult. Bridging this gap may require paradigms that embed learning mechanisms within interactive social contexts. A promising approach to addressing this limitation lies in the integration of eye-tracking and realistic social stimuli in learning tasks to relate internal model updates (e.g. prediction errors) to social attention.

Some studies have taken steps in this direction. For example, Ganglmayer et al., 2020 employed an observational learning task in which participants watched animated agents perform goal-directed actions. The task was designed to infer the agents' hidden goals, allowing for model-based estimation of prediction error signals during action observation. Eye-tracking was used to quantify attention to task-relevant cues (e.g., the agent's hand or goal object), thereby linking gaze patterns to the process of updating beliefs about others' intentions. Autistic participants showed different eye movement patterns and altered learning dynamics, suggesting alterations in how social prediction errors are registered and used to guide future expectations. Similarly, Krogh-Jespersen et al., 2018 used an action-learning paradigm in infants and young children, where participants observed human agents reaching for objects in a naturalistic setting. Eye-tracking data were analyzed in relation to the predictability and informativeness of the observed actions, providing an implicit measure of learning. Although computational modeling was not applied in full, gaze behavior served as a proxy for updating social expectations, suggesting a framework through which attention and learning could be coupled more tightly.

Together, these studies illustrate how realistic social stimuli and gaze behavior can be combined to operationalize learning processes in a way that better approximates real-world social cognition. Such studies represent a promising approach toward capturing how internal prediction mechanisms may support social attention in naturalistic interactions.

Pharmacology and Physiological Measures

Several limitations should further be noted regarding the pharmacological design. Most importantly, the oxytocin dose used in this study (10 IU via nebulizer) was lower than the conventional 24 IU frequently used in intranasal oxytocin studies. This choice was motivated by dose-response studies, which reported greater efficacy of lower oxytocin doses in modulating social cognitive and attentional processes. Specifically, Quintana et al., 2015, 2016 demonstrated behavioral and neural effects at doses as low as 8 IU. Additionally, higher doses of oxytocin may increase the likelihood of cross-activation of vasopressin receptors, particularly under conditions of opioid receptor blockade. Nevertheless, it remains uncertain whether 10 IU was sufficient to influence naturalistic social gaze in the current setup. A dose-response study investigating effects on social behavior under naturalistic conditions could help determine the optimal dosage for our design.

Furthermore, the route of administration differed between the two compounds, with oxytocin delivered intranasally and naltrexone administered orally. Animal studies examining combined effects on social attention employed intranasal delivery of both oxytocin and *naloxone* (an opioid antagonist delivered intranasally) thereby harmonizing administration routes and potentially enhancing simultaneous central bioavailability (dal monte study). Translating this approach to human research may offer methodological advantages by minimizing confounds due to route-specific differences in absorption or onset of action. Future studies could consider simultaneous intranasal delivery of naloxone and oxytocin to facilitate direct comparisons and improve interpretability of combined pharmacological effects.

Additionally, pharmacodynamic responses likely vary as a function of genetic variability (e.g., OXTR or OPRM1 polymorphisms), which were not assessed in this study. Variants of receptor-coding genes, such as the A allele of OXTR, as well as the G allele of OPRM1 (A118G), have been associated with altered social behavior and ASD (LoParo & Waldman, 2015; Troisi et al., 2011). Future studies could account for this by genotyping (e.g. from saliva samples) to identify subgroups in neurotypical and ASD participants that may differ in their responsiveness to oxytocin or opioid antagonists.

Regarding the measurement of HPA axis activity, the number of sampling points was limited, which restricted the ability to capture dynamic fluctuations in stress reactivity. Future studies should include more frequent cortisol sampling to better characterize temporal patterns of hormonal response. With more sampling points, measures such as the area under the curve with respect to increase (AUCi) or the MinMax approach could provide a more robust assessment of cortisol reactivity. Additionally, collecting a later baseline measure could help ensure that participants are in a relaxed state, as they may experience elevated arousal upon arrival at the lab. The absence of an explicit stress induction paradigm further limits the interpretability of cortisol levels and the context in which pharmacological effects occurred. Future studies could examine how acute social stress, i.e. via the TSST, modulates gaze behavior and interacts with pharmacological interventions.

5 Conclusion

This thesis set out with the premise that sociality is deeply embedded in the central nervous system and that targeted neurobiological manipulation can help to uncover the mechanisms which govern it. I aimed to address this by drawing on several lines of research, including reward processing and motivation, the common currency hypothesis, the social motivation hypothesis, as well as predictive coding, computational modeling and reinforcement learning, alongside state-dependent models of opioidergic effects. A central goal was to advance the methodological toolkit of social neuroscience by developing naturalistic paradigms that move beyond artificial, screen-based settings which may fail to capture the complexities of real-life social behavior.

A major finding of this thesis is that social gaze in ASD differs reliably from neurotypical controls during real-life social interaction, and that these differences can be robustly measured using naturalistic paradigms. This underscores the importance of grounding social neuroscience in methods that test the predictions of neurobiological models within realistic contexts. Without such efforts, theories built on artificial tasks and highly controlled laboratory settings risk limited generalizability to real-world social behavior, ultimately constraining both theoretical progress and translational impact.

Many open questions remain, and new ones have been raised. For instance, the lack of a main effect of oxytocin and opioid does not refute the involvement of reward-related processes in social behavior. Additionally, although findings regarding the social effects of opioids and oxytocin may sometimes have been inconsistent or overstated in the literature, it is unlikely that they exert no influence on social behavior at all. Rather, further extensions and refinements of the paradigms developed here are needed to explore the precise conditions under which these systems modulate social attention and learning, and to determine whether social differences in ASD reflect truly social impairments or more domain-general alterations in reward sensitivity and arousal regulation.

Furthermore, while the findings offer some evidence challenging the social motivation hypothesis, they also highlight the need for further extensions of prediction and learning paradigms. Predictive processing is often operationalized in very artificial experimental designs and the implications for real-life social differences observed in ASD are questionable. In general, within the predictive processing framework, it would be important to move beyond narrow operationalizations of prediction and learning and instead connect theoretical constructs to observable social behavior with clinical relevance. While this thesis sought to bridge the gap between predictive coding and naturalistic social attention, two domains with profound mutual implications that are rarely examined together, fully integrating eye tracking as a measure of learning and prediction in realistic social contexts would be necessary to truly connect these domains.

Lastly, the overarching goal of ASD research should not be limited to identifying the causes of social differences, but must extend to improving the lives of individuals who face social challenges. While uncovering the mechanisms underlying social differences is important to address where support is needed and how it can be provided, many people within the autistic community advocate for a shift in focus: away from pathologizing differences, and toward understanding how social environments can be adapted to support diverse social needs. Variability in social behavior is a natural and expected part of a continuous spectrum that encompasses the full range of human experience, one that follows a normal distribution rather than a binary divide. Furthermore, science itself is a fundamentally social activity, and its goals are shaped largely by the societies in which it is embedded. Understanding and supporting diverse social needs is therefore not only a scientific task but itself a social one. This thesis aimed to have made a modest contribution to that collective effort.

6 References

Alvares, G. A., Quintana, D. S., & Whitehouse, A. J. O. (2017). Beyond the hype and hope: Critical considerations for intranasal oxytocin research in autism spectrum disorder. Autism Research, 10(1), 25–41. https://doi.org/10.1002/aur.1692

American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders: DSM-5 (Vol. 5). American psychiatric association Washington, DC.

Andari, E., Duhamel, J.-R., Zalla, T., Herbrecht, E., Leboyer, M., & Sirigu, A. (2010). Promoting social behavior with oxytocin in high-functioning autism spectrum disorders. Proceedings of the National Academy of Sciences, 107(9), 4389–4394. https://doi.org/10.1073/pnas.0910249107

Auyeung, B., Lombardo, M. V., Heinrichs, M., Chakrabarti, B., Sule, A., Deakin, J. B., Bethlehem, R. A. I., Dickens, L., Mooney, N., Sipple, J. A. N., Thiemann, P., & Baron-Cohen, S. (2015). Oxytocin increases eye contact during a real-time, naturalistic social interaction in males with and without autism. Translational Psychiatry, 5(2). https://doi.org/10.1038/tp.2014.146

Bakos, J., Srancikova, A., Havranek, T., & Bacova, Z. (2018). Molecular Mechanisms of Oxytocin Signaling at the Synaptic Connection. Neural Plasticity, 2018(1), 4864107. https://doi.org/10.1155/2018/4864107

Bartz, J. A., Zaki, J., Bolger, N., Hollander, E., Ludwig, N. N., Kolevzon, A., & Ochsner, K. N. (2010). Oxytocin Selectively Improves Empathic Accuracy. Psychological Science, 21(10), 1426–1428. https://doi.org/10.1177/0956797610383439

Beck, A. T., Ward, C., Mendelson, M., Mock, J., & Erbauch, J. (1978). Beck depression inventory (BDI). San Antonio, TX.

Beery, A. K., & Kaufer, D. (2015). Stress, social behavior, and resilience: Insights from rodents. Neurobiology of Stress, 1, 116–127. https://doi.org/10.1016/j.ynstr.2014.10.004

Berridge, K. C., & Kringelbach, M. L. (2015). Pleasure Systems in the Brain. Neuron, 86(3), 646–664. https://doi.org/10.1016/j.neuron.2015.02.018

Berridge, K. C., Robinson, T. E., & Aldridge, J. W. (2009). Dissecting components of reward: 'Liking', 'wanting', and learning. Current Opinion in Pharmacology, 9(1), 65–73. https://doi.org/10.1016/j.coph.2008.12.014

Bicknell, R. J., & Leng, G. (1982). Endogenous opiates regulate oxytocin but not vasopressin secretion from the neurohypophysis. Nature, 298(5870), 161–162.

Bicknell, R. J., Leng, G., Lincoln, D. W., & Russell, J. A. (1988). Naloxone excites oxytocin neurones in the supraoptic nucleus of lactating rats after chronic morphine treatment. The Journal of Physiology, 396(1), 297–317. https://doi.org/10.1113/jphysiol.1988.sp016963

Blake, R., Turner, L. M., Smoski, M. J., Pozdol, S. L., & Stone, W. L. (2003). Visual Recognition of Biological Motion is Impaired in Children With Autism. Psychological Science, 14(2), 151–157. https://doi.org/10.1111/1467-9280.01434

Bouvard, M. P., Leboyer, M., Launay, J.-M., Recasens, C., Plumet, M.-H., Waller-Perotte, D., Tabuteau, F., Bondoux, D., Dugas, M., Lensing, P., & Panksepp, J. (1995). Low-dose naltrexone effects on plasma chemistries and clinical symptoms in autism: A double-blind, placebo-controlled study. Psychiatry Research, 58(3), 191–201. https://doi.org/10.1016/0165-1781(95)02601-R

Brown, C. H., Russell, J. A., & Leng, G. (2000). Opioid modulation of magnocellular neurosecretory cell activity. Neuroscience Research, 36(2), 97–120. https://doi.org/10.1016/S0168-0102(99)00121-2

Brownstein, M. J. (1993). A brief history of opiates, opioid peptides, and opioid receptors. Proceedings of the National Academy of Sciences, 90(12), 5391–5393. https://doi.org/10.1073/pnas.90.12.5391

Bruchas, M. R., Land, B. B., & Chavkin, C. (2010). The dynorphin/kappa opioid system as a modulator of stress-induced and pro-addictive behaviors. Brain Research, 1314, 44–55. https://doi.org/10.1016/j.brainres.2009.08.062

Buehler, R., Silani, G., & Ansorge, U. (2024). Social attention in the wild—Fixations to the eyes and autistic traits during a naturalistic interaction in a healthy sample. https://doi.org/10.21203/rs.3.rs-4016436/v1

Burkett, J. P., Spiegel, L. L., Inoue, K., Murphy, A. Z., & Young, L. J. (2011). Activation of μ-Opioid Receptors in the Dorsal Striatum is Necessary for Adult Social Attachment in Monogamous Prairie Voles. Neuropsychopharmacology, 36(11), 2200–2210. https://doi.org/10.1038/npp.2011.117

Cañigueral, R., & Hamilton, A. F. de C. (2019). The role of eye gaze during natural social interactions in typical and autistic people. Frontiers in Psychology, 10, 560.

Cañigueral, R., Hamilton, A., & Ward, J. A. (2018). Don't Look at Me, I'm Wearing an Eyetracker! Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers, 994–998. https://doi.org/10.1145/3267305.3274123

Cañigueral, R., Ward, J. A., & Hamilton, A. F. de C. (2021). Effects of being watched on eye gaze and facial displays of typical and autistic individuals during conversation. Autism, 25(1), 210–226. https://doi.org/10.1177/1362361320951691

Cannon, J., O'Brien, A. M., Bungert, L., & Sinha, P. (2021). Prediction in Autism Spectrum Disorder: A Systematic Review of Empirical Evidence. Autism Research, 14(4), 604–630. https://doi.org/10.1002/aur.2482

Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E., & Sheikh, Y. (2021). OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(1), 172–186. IEEE Transactions on Pattern Analysis and Machine Intelligence. https://doi.org/10.1109/TPAMI.2019.2929257

Cardoso, C., Ellenbogen, M. A., Orlando, M. A., Bacon, S. L., & Joober, R. (2013). Intranasal oxytocin attenuates the cortisol response to physical stress: A dose–response study. Psychoneuroendocrinology, 38(3), 399–407. https://doi.org/10.1016/j.psyneuen.2012.07.013

Cardoso, C., Kingdon, D., & Ellenbogen, M. A. (2014). A meta-analytic review of the impact of intranasal oxytocin administration on cortisol concentrations during laboratory tasks: Moderation by method and mental health. Psychoneuroendocrinology, 49(1), 161–170. https://doi.org/10.1016/j.psyneuen.2014.07.014

Chang, S. W. C., Barter, J. W., Ebitz, R. B., Watson, K. K., & Platt, M. L. (2012). Inhaled oxytocin amplifies both vicarious reinforcement and self reinforcement in rhesus macaques (Macaca mulatta). Proceedings of the National Academy of Sciences of the United States of America, 109(3), 959–964. https://doi.org/10.1073/pnas.1114621109

Chelnokova, O., Laeng, B., Eikemo, M., Riegels, J., Løseth, G., Maurud, H., Willoch, F., & Leknes, S. (2014). Rewards of beauty: The opioid system mediates social motivation in humans. Molecular Psychiatry, 19(7), 746–747.

Chelnokova, O., Laeng, B., Løseth, G., Eikemo, M., Willoch, F., & Leknes, S. (2016). The μ-opioid system promotes visual attention to faces and eyes. Social Cognitive and Affective Neuroscience, 11(12), 1902–1909. https://doi.org/10.1093/scan/nsw116

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism.

Trends in Cognitive Sciences, 16(4), 231–239. https://doi.org/10.1016/j.tics.2012.02.007

Chevallier, C., Parish-Morris, J., Mcvey, A., Rump, K. M., Sasson, N. J., Herrington, J. D., & Schultz, R. T. (2015). Measuring social attention and motivation in autism spectrum disorder using eye-tracking: Stimulus type matters. Autism Research, 8(5), 620–628. https://doi.org/10.1002/aur.1479

Chita-Tegmark, M. (2016). Social attention in ASD: A review and meta-analysis of eye-tracking studies. Research in Developmental Disabilities, 48, 79–93. https://doi.org/10.1016/j.ridd.2015.10.011

Chu Sin Chung, P., & Kieffer, B. L. (2013). Delta opioid receptors in brain function and diseases. Pharmacology & Therapeutics, 140(1), 112–120. https://doi.org/10.1016/j.pharmthera.2013.06.003

Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. British Journal of Clinical Psychology, 43(3), 245–265. https://doi.org/10.1348/0144665031752934

Cuve, H. C., Gao, Y., & Fuse, A. (2018). Is it avoidance or hypoarousal? A systematic review of emotion recognition, eye-tracking, and psychophysiological studies in young adults with autism spectrum conditions. Research in Autism Spectrum Disorders, 55, 1–13. https://doi.org/10.1016/j.rasd.2018.07.002

Dal Monte, O., Piva, M., Anderson, K. M., Tringides, M., Holmes, A. J., & Chang, S. W. C. (2017). Oxytocin under opioid antagonism leads to supralinear enhancement of social attention. Proceedings of the National Academy of Sciences of the United States of America, 114(20), 5247–5252. https://doi.org/10.1073/pnas.1702725114

Dal Monte, O., Piva, M., Morris, J. A., & Chang, S. W. C. (2016). Live interaction distinctively shapes social gaze dynamics in rhesus macaques. Journal of Neurophysiology, 116(4), 1626–1643. https://doi.org/10.1152/jn.00442.2016

Davis, M. H. (1980). Interpersonal reactivity index. https://psycnet.apa.org/doiLanding?doi=10.1037/t01093-000

Di Simplicio, M., Massey-Chase, R., Cowen, P., & Harmer, C. (2009). Oxytocin enhances processing of positive versus negative emotional information in healthy male volunteers. Journal of Psychopharmacology, 23(3), 241–248. https://doi.org/10.1177/0269881108095705

du Vigneaud, V., Ressler, C., Swan, J. M., Roberts, C. W., & Katsoyannis, P. G. (1954). The Synthesis of Oxytocin. Journal of the American Chemical Society, 76(12), 3115–3121. https://doi.org/10.1021/ja01641a004

Dunbar, R. I. M. (1998). The social brain hypothesis. Evolutionary Anthropology: Issues, News, and Reviews, 6(5), 178–190. https://doi.org/10.1002/(SICI)1520-6505(1998)6:5<178::AID-EVAN5>3.0.CO;2-8

Ebitz, R. B., Watson, K. K., & Platt, M. L. (2013). Oxytocin blunts social vigilance in the rhesus macaque. Proceedings of the National Academy of Sciences, 110(28), 11630–11635. https://doi.org/10.1073/pnas.1305230110

Eisenberger, N. I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. Nature Reviews Neuroscience, 13(6), 421–434.

Ellingsen, D.-M., Leknes, S., Løseth, G., Wessberg, J., & Olausson, H. (2016). The Neurobiology Shaping Affective Touch: Expectation, Motivation, and Meaning in the Multisensory Context. Frontiers in Psychology, 6. https://doi.org/10.3389/fpsyg.2015.01986

Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. Neuroscience & Biobehavioral Reviews, 24(6), 581–604. https://doi.org/10.1016/S0149-7634(00)00025-7

Engel, S., Klusmann, H., Ditzen, B., Knaevelsrud, C., & Schumacher, S. (2019). Menstrual cycle-related fluctuations in oxytocin concentrations: A systematic review and meta-analysis. Frontiers in Neuroendocrinology, 52(August 2018), 144–155. https://doi.org/10.1016/j.yfrne.2018.11.002

Ewbank, M. P., Pell, P. J., Powell, T. E., von dem Hagen, E. A. H., Baron-Cohen, S., & Calder, A. J. (2017). Repetition Suppression and Memory for Faces is Reduced in Adults with Autism Spectrum Conditions. Cerebral Cortex, 27(1), 92–103. https://doi.org/10.1093/cercor/bhw373

Fabre-Nys, C., Meller, R. E., & Keverne, E. B. (1982). Opiate antagonists stimulate affiliative behaviour in monkeys. Pharmacology Biochemistry and Behavior, 16(4), 653–659. https://doi.org/10.1016/0091-3057(82)90432-4

Fan, S., Weinberg-Wolf, H., Piva, M., Dal Monte, O., & Chang, S. W. C. (2020). Combinatorial Oxytocin Neuropharmacology in Social Cognition. Trends in Cognitive Sciences, 24(1), 8–12.

Frazier, T. W., Strauss, M., Klingemier, E. W., Zetzer, E. E., Hardan, A. Y., Eng, C., & Youngstrom, E. A. (2017). A Meta-Analysis of Gaze Differences to Social and Nonsocial Information Between Individuals With and Without Autism. Journal of the American Academy of Child & Adolescent Psychiatry, 56(7), 546–555. https://doi.org/10.1016/j.jaac.2017.05.005

Freeth, M., Foulsham, T., & Kingstone, A. (2013). What Affects Social Attention? Social Presence, Eye Contact and Autistic Traits. PLOS ONE, 8(1), e53286. https://doi.org/10.1371/journal.pone.0053286

Freitag, C. M., Konrad, C., Häberlen, M., Kleser, C., von Gontard, A., Reith, W., Troje, N. F., & Krick, C. (2008). Perception of biological motion in autism spectrum disorders. Neuropsychologia, 46(5), 1480–1494. https://doi.org/10.1016/j.neuropsychologia.2007.12.025

Freitag, C. M., Retz-Junginger, P., Retz, W., Seitz, C., Palmason, H., Meyer, J., Rösler, M., & von Gontard, A. (2007). Evaluation der deutschen Version des Autismus-Spektrum-Quotienten (AQ)—Die Kurzversion AQ-k. Zeitschrift Für Klinische Psychologie Und Psychotherapie, 36(4), 280–289. https://doi.org/10.1026/1616-3443.36.4.280

Frith, C. D. (2007). The social brain? Philosophical Transactions of the Royal Society B: Biological Sciences, 362(1480), 671–678. https://doi.org/10.1098/rstb.2006.2003

Gillberg, C., Terenius, L., Hagberg, B., Witt-Engerström, I., & Eriksson, I. (1990). CSF beta-endorphins in childhood neuropsychiatric disorders. Brain and Development, 12(1), 88–92.

Gonzalez, J. P., & Brogden, R. N. (1988). Naltrexone: A Review of its Pharmacodynamic and Pharmacokinetic Properties and Therapeutic Efficacy in the Management of Opioid Dependence. Drugs, 35(3), 192–213. https://doi.org/10.2165/00003495-198835030-00002

Gossen, A., Hahn, A., Westphal, L., Prinz, S., Schultz, R. T., Gründer, G., & Spreckelmeyer, K. N. (2012). Oxytocin plasma concentrations after single intranasal oxytocin administration – A study in healthy men. Neuropeptides, 46(5), 211–215. https://doi.org/10.1016/j.npep.2012.07.001

Gregory, N. J., Bolderston ,Helen, & and Antolin, J. V. (2019). Attention to faces and gaze-following in social anxiety: Preliminary evidence from a naturalistic eye-tracking investigation. Cognition and Emotion, 33(5), 931–942. https://doi.org/10.1080/02699931.2018.1519497

Gregory, S. G., Connelly, J. J., Towers, A. J., Johnson, J., Biscocho, D., Markunas, C. A., Lintas, C., Abramson, R. K., Wright, H. H., Ellis, P., Langford, C. F., Worley, G., Delong, G. R., Murphy, S. K., Cuccaro, M. L., Persico, A., & Pericak-Vance, M. A. (2009).

Genomic and epigenetic evidence for oxytocin receptor deficiency in autism. BMC Medicine, 7(1), Article 1. https://doi.org/10.1186/1741-7015-7-62

Grinevich, V., & Neumann, I. D. (2021). Brain oxytocin: How puzzle stones from animal studies translate into psychiatry. Molecular Psychiatry, 26(1), 265–279. https://doi.org/10.1038/s41380-020-0802-9

Grippo, A. J., Trahanas, D. M., Zimmerman, R. R., Porges, S. W., & Carter, C. S. (2009). Oxytocin protects against negative behavioral and autonomic consequences of long-term social isolation. Psychoneuroendocrinology, 34(10), 1542–1553. https://doi.org/10.1016/j.psyneuen.2009.05.017

Grzadzinski, R., Huerta, M., & Lord, C. (2013). DSM-5 and autism spectrum disorders (ASDs): An opportunity for identifying ASD subtypes. Molecular Autism, 4(1), Article 1. https://doi.org/10.1186/2040-2392-4-12

Guastella, A. J., Einfeld, S. L., Gray, K. M., Rinehart, N. J., Tonge, B. J., Lambert, T. J., & Hickie, I. B. (2010). Intranasal Oxytocin Improves Emotion Recognition for Youth with Autism Spectrum Disorders. Biological Psychiatry, 67(7), 692–694. https://doi.org/10.1016/j.biopsych.2009.09.020

Guastella, A. J., Gray, K. M., Rinehart, N. J., Alvares, G. A., Tonge, B. J., Hickie, I. B., Keating, C. M., Cacciotti-Saija, C., & Einfeld, S. L. (2015). The effects of a course of intranasal oxytocin on social behaviors in youth diagnosed with autism spectrum disorders: A randomized controlled trial. Journal of Child Psychology and Psychiatry, 56(4), 444–452. https://doi.org/10.1111/jcpp.12305

Guastella, A. J., & Hickie, I. B. (2016). Oxytocin treatment, circuitry, and autism: A critical review of the literature placing oxytocin into the autism context. Biological Psychiatry, 79(3), 234–242.

Guastella, A. J., Mitchell, P. B., & Dadds, M. R. (2008). Oxytocin Increases Gaze to the Eye Region of Human Faces. Biological Psychiatry, 63(1), 3–5. https://doi.org/10.1016/j.biopsych.2007.06.026

Günther, V., Kropidlowski, A., Schmidt, F. M., Koelkebeck, K., Kersting, A., & Suslow, T. (2021). Attentional processes during emotional face perception in social anxiety disorder: A systematic review and meta-analysis of eye-tracking findings. Progress in Neuro-Psychopharmacology and Biological Psychiatry, 111, 110353. https://doi.org/10.1016/j.pnpbp.2021.110353

Heimberg, R. G., Mueller, G. P., Holt, C. S., Hope, D. A., & Liebowitz, M. R. (1992). Assessment of anxiety in social interaction and being observed by others: The social interaction anxiety scale and the Social Phobia Scale. Behavior Therapy, 23(1), 53–73. https://doi.org/10.1016/S0005-7894(05)80308-9

Heinrichs, M., Baumgartner, T., Kirschbaum, C., & Ehlert, U. (2003). Social support and oxytocin interact to suppress cortisol and subjective responses to psychosocial stress. Biological Psychiatry, 54(12), 1389–1398. https://doi.org/10.1016/S0006-3223(03)00465-7

Hermesch, A. C., Kernberg, A. S., Layoun, V. R., & Caughey, A. B. (2024). Oxytocin: Physiology, pharmacology, and clinical application for labor management. American Journal of Obstetrics and Gynecology, 230(3, Supplement), S729–S739. https://doi.org/10.1016/j.ajog.2023.06.041

Hu, J., Qi, S., Becker, B., Luo, L., Gao, S., Gong, Q., Hurlemann, R., & Kendrick, K. M. (2015). Oxytocin selectively facilitates learning with social feedback and increases activity and functional connectivity in emotional memory and reward processing regions. Human Brain Mapping, 36(6), 2132–2146. https://doi.org/10.1002/hbm.22760

Hubert, B., Wicker, B., Moore, D. G., Monfardini, E., Duverger, H., Fonséca, D. D., & Deruelle, C. (2007). Brief Report: Recognition of Emotional and Non-emotional Biological Motion in Individuals with Autistic Spectrum Disorders. Journal of Autism and Developmental Disorders, 37(7), 1386–1392. https://doi.org/10.1007/s10803-006-0275-y

Hurlemann, R., Patin, A., Onur, O. A., Cohen, M. X., Baumgartner, T., Metzler, S., Dziobek, I., Gallinat, J., Wagner, M., Maier, W., & Kendrick, K. M. (2010). Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans. Journal of Neuroscience, 30(14), 4999–5007. https://doi.org/10.1523/JNEUROSCI.5538-09.2010

Huys, Q. J., Pizzagalli, D. A., Bogdan, R., & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: A behavioural meta-analysis. Biology of Mood & Anxiety Disorders, 3(1), Article 1. https://doi.org/10.1186/2045-5380-3-12

Inagaki, T. K., Ray, L. A., Irwin, M. R., Way, B. M., & Eisenberger, N. I. (2016). Opioids and social bonding: Naltrexone reduces feelings of social connection. Social Cognitive and Affective Neuroscience, 11(5), 728–735. https://doi.org/10.1093/scan/nsw006

Insel, T. R., & Shapiro, L. E. (1992). Oxytocin receptor distribution reflects social organization in monogamous and polygamous voles. Proceedings of the National Academy of Sciences, 89(13), 5981–5985. https://doi.org/10.1073/pnas.89.13.5981

Jongerius, C., Callemein, T., Goedemé, T., Van Beeck, K., Romijn, J. A., Smets, E. M. A., & Hillen, M. A. (2021). Eye-tracking glasses in face-to-face interactions: Manual versus automated assessment of areas-of-interest. Behavior Research Methods, 53(5), 2037–2048. https://doi.org/10.3758/s13428-021-01544-2

Joseph, R. M., Tager-Flusberg, H., & Lord, C. (2002). Cognitive profiles and social-communicative functioning in children with autism spectrum disorder. Journal of Child Psychology and Psychiatry, 43(6), 807–821. https://doi.org/10.1111/1469-7610.00092

Kalin, N. H., Shelton, S. E., & Barksdale, C. M. (1988). Opiate modulation of separation-induced distress in non-human primates. Brain Research, 440(2), 285–292. https://doi.org/10.1016/0006-8993(88)90997-3

Kanat, M., Heinrichs, M., & Domes, G. (2014). Oxytocin and the social brain: Neural mechanisms and perspectives in human research. Brain Research, 1580, 160–171. https://doi.org/10.1016/j.brainres.2013.11.003

Kassner, M., Patera, W., & Bulling, A. (2014). Pupil. Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, 1151–1160. https://doi.org/10.1145/2638728.2641695

Keating, C. T., Fraser, D. S., Sowden, S., & Cook, J. L. (2022). Differences between autistic and non-autistic adults in the recognition of anger from facial motion remain after controlling for alexithymia. Journal of Autism and Developmental Disorders, 52(4), 1855–1871.

Kendrick, K. M., Guastella, A. J., & Becker, B. (2017). Overview of Human Oxytocin Research. In Behavioral Pharmacology of Neuropeptides: Oxytocin (pp. 321–348). Springer, Cham. https://doi.org/10.1007/7854_2017_19

Keverne, E. B., Martensz, N. D., & Tuite, B. (1989). Beta-endorphin concentrations in cerebrospinal fluid of monkeys are influenced by grooming relationships. Psychoneuroendocrinology, 14(1), 155–161. https://doi.org/10.1016/0306-4530(89)90065-6

Kinard, J. L., Mosner, M. G., Greene, R. K., Addicott, M., Bizzell, J., Petty, C., Cernasov, P., Walsh, E., Eisenlohr-Moul, T., Carter, R. M., McLamb, M., Hopper, A., Sukhu, R., & Dichter, G. S. (2020). Neural Mechanisms of Social and Nonsocial Reward Prediction Errors in Adolescents with Autism Spectrum Disorder. Autism Research, 13(5), 715–728. https://doi.org/10.1002/aur.2273

King, A. (2002). Hypothalamic-Pituitary-Adrenocortical (HPA) Axis Response and Biotransformation of Oral Naltrexone Preliminary Examination of Relationship to Family History of Alcoholism.

Neuropsychopharmacology, 26(6), 778–788. https://doi.org/10.1016/S0893-133X(01)00416-X

Kirsch, P., Esslinger, C., Chen, Q., Mier, D., Lis, S., Siddhanti, S., Gruppe, H., Mattay, V. S., Gallhofer, B., & Meyer-Lindenberg, A. (2005). Oxytocin modulates neural circuitry for social cognition and fear in humans. Journal of Neuroscience, 25(49), 11489–11493.

Knoll, A. T., & Carlezon, W. A. (2010). Dynorphin, stress, and depression. Brain Research, 1314, 56–73. https://doi.org/10.1016/j.brainres.2009.09.074

Knoll, A. T., Muschamp, J. W., Daws, S. E., Ferguson, D., Dietz, D. M., Meloni, E. G., Carroll, F. I., Nestler, E. J., Konradi, C., & Carlezon, W. A. (2011). Kappa Opioid Receptor Signaling in the Basolateral Amygdala Regulates Conditioned Fear and Anxiety in Rats. Biological Psychiatry, 70(5), 425–433. https://doi.org/10.1016/j.biopsych.2011.03.017

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. Nature, 435(7042), 673–676. https://doi.org/10.1038/nature03701

Kudielka, B. M., Hellhammer, D. H., & Kirschbaum, C. (2007). Ten years of research with the Trier Social Stress Test (TSST)-revisited. Social Neuroscience: Integrating Biological and Psychological Explanations of Social Behavior, 56–83.

Kumsta, R., & Heinrichs, M. (2013). Oxytocin, stress and social behavior: Neurogenetics of the human oxytocin system. Current Opinion in Neurobiology, 23(1), 11–16. https://doi.org/10.1016/j.conb.2012.09.004

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. Journal of Statistical Software, 82(13). https://doi.org/10.18637/jss.v082.i13

Lalanne, L., Ayranci, G., Filliol, D., Gavériaux-Ruff, C., Befort, K., Kieffer, B. L., & Lutz, P.-E. (2017). Kappa opioid receptor antagonism and chronic antidepressant treatment have beneficial activities on social interactions and grooming deficits during heroin abstinence. Addiction Biology, 22(4), 1010–1021. https://doi.org/10.1111/adb.12392

Land, M., Mennie, N., & Rusted, J. (1999). The Roles of Vision and Eye Movements in the Control of Activities of Daily Living. Perception, 28(11), 1311–1328. https://doi.org/10.1068/p2935

Lawson, R. P., Mathys, C., & Rees, G. (2017). Adults with autism overestimate the volatility of the sensory environment. Nature Neuroscience, 20(9), 1293–1299. https://doi.org/10.1038/nn.4615

Lawson, R. P., Rees, G., & Friston, K. J. (2014). An aberrant precision account of autism. Frontiers in Human Neuroscience, 8. https://doi.org/10.3389/fnhum.2014.00302

Leboyer, M., Bouvard, M. P., Launay, J.-M., Tabuteau, F., Waller, D., Dugas, M., Kerdelhue, B., Lensing, P., & Panksepp, J. (1992). Brief report: A double-blind study of naltrexone in infantile autism. Journal of Autism and Developmental Disorders, 22(2), 309–319. https://doi.org/10.1007/BF01058158

Leboyer, M., Philippe, A., Bouvard, M., Guilloud-Bataille, M., Bondoux, D., Tabuteau, F., Feingold, J., Mouren-Simeoni, M.-C., & Launay, J.-M. (1999). Whole blood serotonin and plasma beta-endorphin in autistic probands and their first-degree relatives. Biological Psychiatry, 45(2), 158–163. https://doi.org/10.1016/S0006-3223(97)00532-5

Lehrl, S. (1999). Mehrfachwahl-Wortschatz-Intelligenztest: MWT-B. Spitta.

Lenth, R. (2022). emmeans: Estimated marginal means, aka least-squares means. R package version 1.7. 2.

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: A neural common currency for choice. Current Opinion in Neurobiology, 22(6), 1027–1038.

Lieder, I., Adam, V., Frenkel, O., Jaffe-Dax, S., Sahani, M., & Ahissar, M. (2019). Perceptual bias reveals slow-updating in autism and fast-forgetting in dyslexia. Nature Neuroscience, 22(2), 256–264. https://doi.org/10.1038/s41593-018-0308-9

Liu, Y., & Wang, Z. X. (2003). Nucleus accumbens oxytocin and dopamine interact to regulate pair bond formation in female prairie voles. Neuroscience, 121(3), 537–544. https://doi.org/10.1016/S0306-4522(03)00555-4

LoParo, D., & Waldman, I. D. (2015). The oxytocin receptor gene (OXTR) is associated with autism spectrum disorder: A meta-analysis. Molecular Psychiatry, 20(5), 640–646.

Løseth, G. E., Ellingsen, D.-M., & Leknes, S. (2014). State-dependent μ-opioid modulation of social motivation. Frontiers in Behavioral Neuroscience, 8, 430.

Loseth, G. E., Ellingsen, D.-M., & Leknes, S. (2014). State-dependent μ-opioid Modulation of Social Motivation–a model. Frontiers in Behavioral Neuroscience, 8, 430.

Løseth, G., Trøstheim, M., & Leknes, S. (2024). Endogenous mu-opioid modulation of social connection in humans: A systematic review and meta-analysis. Translational Psychiatry, 14(1), 1–11. https://doi.org/10.1038/s41398-024-03088-3

Lukas, M., Toth, I., Reber, S. O., Slattery, D. A., Veenema, A. H., & Neumann, I. D. (2011). The Neuropeptide Oxytocin Facilitates Pro-Social Behavior and Prevents Social Avoidance in Rats and Mice. Neuropsychopharmacology, 36(11), 2159–2168. https://doi.org/10.1038/npp.2011.95

Macdonald, R. G., & Tatler, B. W. (2018). Gaze in a real-world social interaction: A dual eye-tracking study. Quarterly Journal of Experimental Psychology, 71(10), 2162–2173. https://doi.org/10.1177/1747021817739221

Machin, A. J., & Dunbar, R. I. M. (2011). The brain opioid theory of social attachment: A review of the evidence. https://doi.org/10.1163/000579511X596624

Mahler, S. V., & Berridge, K. C. (2009). Which cue to "want?" Central amygdala opioid activation enhances and focuses incentive salience on a prepotent reward cue. Journal of Neuroscience, 29(20), 6500–6513.

Manduca, A., Carbone, E., Schiavi, S., Cacchione, C., Buzzelli, V., Campolongo, P., & Trezza, V. (2021). The neurochemistry of social reward during development: What have we learned from rodent models? Journal of Neurochemistry, 157(5), 1408–1435. https://doi.org/10.1111/jnc.15321

Manninen, S., Tuominen, L., Dunbar, R. I., Karjalainen, T., Hirvonen, J., Arponen, E., Hari, R., Jääskeläinen, I. P., Sams, M., & Nummenmaa, L. (2017). Social Laughter Triggers Endogenous Opioid Release in Humans. Journal of Neuroscience, 37(25), 6125–6131. https://doi.org/10.1523/JNEUROSCI.0688-16.2017

Martel, F. L., Nevison, C. M., Rayment, F. D., Simpson, M. J. A., & Keverne, E. B. (1993). Opioid receptor blockade reduces maternal affect and social grooming in rhesus monkeys. Psychoneuroendocrinology, 18(4), 307–321. https://doi.org/10.1016/0306-4530(93)90027-I

Martins, D., Brodmann, K., Veronese, M., Dipasquale, O., Mazibuko, N., Schuschnig, U., Zelaya, F., Fotopoulou, A., & Paloyelis, Y. (2022). "Less is more": A dose-response account of intranasal oxytocin pharmacodynamics in the human brain. Progress in Neurobiology, September 2021, 102239. https://doi.org/10.1016/j.pneurobio.2022.102239

Mathys, C., Daunizeau, J., Friston, K. J., & Stephan, K. E. (2011). A Bayesian Foundation for Individual Learning Under Uncertainty. Frontiers in Human Neuroscience, 5. https://doi.org/10.3389/fnhum.2011.00039

Mayrand, F., Capozzi, F., & Ristic, J. (2023). A dual mobile eye tracking study on natural eye contact during live interactions. Scientific Reports, 13(1), 11385.

McGraw, L. A., & Young, L. J. (2010). The prairie vole: An emerging model organism for understanding the social brain. Trends in Neurosciences, 33(2), 103–109. https://doi.org/10.1016/j.tins.2009.11.006

McLaughlin, J. P., Li, S., Valdez, J., Chavkin, T. A., & Chavkin, C. (2006). Social Defeat Stress-Induced Behavioral

Responses are Mediated by the Endogenous Kappa Opioid System. Neuropsychopharmacology, 31(6), 1241–1248. https://doi.org/10.1038/sj.npp.1300872

Miller, R., Wojtyniak, J.-G., Weckesser, L. J., Alexander, N. C., Engert, V., & Lehr, T. (2018). How to disentangle psychobiological stress reactivity and recovery: A comparison of model-based and non-compartmental analyses of cortisol concentrations. Psychoneuroendocrinology, 90, 194–210. https://doi.org/10.1016/j.psyneuen.2017.12.019

Milton, D. E. M. (2012). On the ontological status of autism: The 'double empathy problem.' Disability & Society, 27(6), 883–887. https://doi.org/10.1080/09687599.2012.710008

Modahl, C., Green, L. A., Fein, D., Morris, M., Waterhouse, L., Feinstein, C., & Levin, H. (1998). Plasma oxytocin levels in autistic children. Biological Psychiatry, 43(4), 270–277. https://doi.org/10.1016/S0006-3223(97)00439-3

Moles, A., Kieffer, B. L., & D'Amato, F. R. (2004). Deficit in Attachment Behavior in Mice Lacking the μ-Opioid Receptor Gene. Science, 304(5679), 1983–1986. https://doi.org/10.1126/science.1095943

Moriuchi, J. M., Klin, A., & Jones, W. (2017). Mechanisms of Diminished Attention to Eyes in Autism. American Journal of Psychiatry, 174(1), 26–35. https://doi.org/10.1176/appi.ajp.2016.15091222

Nagamitsu, S., Matsuishi, T., Kisa, T., Komori, H., Miyazaki, M., Hashimoto, T., Yamashita, Y., Ohtaki, E., & Kato, H. (1997). CSF β-Endorphin Levels in Patients with Infantile Autism. Journal of Autism and Developmental Disorders, 27(2), 155–163. https://doi.org/10.1023/A:1025839807431

Neumann, I. D., & Landgraf, R. (2012). Balance of brain oxytocin and vasopressin: Implications for anxiety, depression, and social behaviors. Trends in Neurosciences, 35(11), 649–659. https://doi.org/10.1016/j.tins.2012.08.004

Nisbett, K. E., Vendruscolo, L. F., & Koob, G. F. (2024). M-Opioid receptor antagonism facilitates the anxiolytic-like effect of oxytocin in mice. Translational Psychiatry, 14(1), 1–12. https://doi.org/10.1038/s41398-024-02830-1

Noris, B., Nadel, J., Barker, M., Hadjikhani, N., & Billard, A. (2012). Investigating gaze of children with ASD in naturalistic settings.

Nummenmaa, L., & Calder, A. J. (2009). Neural mechanisms of social attention. Trends in Cognitive Sciences, 13(3), 135–143. https://doi.org/10.1016/j.tics.2008.12.006

Nummenmaa, L., Tuominen, L., Dunbar, R., Hirvonen, J., Manninen, S., Arponen, E., Machin, A., Hari, R., Jääskeläinen, I. P., & Sams, M. (2016). Social touch modulates endogenous μ-opioid system activity in humans. NeuroImage, 138, 242–247. https://doi.org/10.1016/j.neuroimage.2016.05.063

Nyström, M., Hooge, I. T. C., Hessels, R. S., Andersson, R., Hansen, D. W., Johansson, R., & Niehorster, D. C. (2025). The fundamentals of eye tracking part 3: How to choose an eye tracker. Behavior Research Methods, 57(2), 1–28. https://doi.org/10.3758/s13428-024-02587-x

Ooi, Y. P., Weng, S.-J., Kossowsky, J., Gerger, H., & Sung, M. (2016). Oxytocin and Autism Spectrum Disorders: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. Pharmacopsychiatry, 50, 5–13. https://doi.org/10.1055/s-0042-109400

Panksepp, J. (1979). A neurochemical theory of autism. Trends in Neurosciences, 2, 174–177.

Parker, K. J., Oztan, O., Libove, R. A., Sumiyoshi, R. D., Jackson, L. P., Karhson, D. S., Summers, J. E., Hinman, K. E., Motonaga, K. S., Phillips, J. M., Carson, D. S., Garner, J. P., & Hardan, A. Y. (2017). Intranasal oxytocin treatment for social deficits and biomarkers of response in children with autism. Proceedings of the National Academy of Sciences, 114(30), 8119–8124. https://doi.org/10.1073/pnas.1705521114

Parr, L. A., Modi, M., Siebert, E., & Young, L. J. (2013). Intranasal oxytocin selectively attenuates rhesus monkeys' attention to negative facial expressions. Psychoneuroendocrinology, 38(9), 1748–1756. https://doi.org/10.1016/j.psyneuen.2013.02.011

Pavlova, M. A. (2012). Biological Motion Processing as a Hallmark of Social Cognition. Cerebral Cortex, 22(5), 981–995. https://doi.org/10.1093/cercor/bhr156

Pellicano, E., & Burr, D. (2012). When the world becomes 'too real': A Bayesian explanation of autistic perception. Trends in Cognitive Sciences, 16(10), 504–510. https://doi.org/10.1016/j.tics.2012.08.009

Pellissier, L. P., Gandía, J., Laboute, T., Becker, J. A. J., & Le Merrer, J. (2018). μ opioid receptor, social behaviour and autism spectrum disorder: Reward matters. British Journal of Pharmacology, 175(14), 2750–2769. https://doi.org/10.1111/bph.13808

Perisic, M., Woolcock, K., Hering, A., Mendel, H., & Muttenthaler, M. (2024). Oxytocin and vasopressin signaling in health and disease. Trends in Biochemical Sciences, 49(4), 361–377. https://doi.org/10.1016/j.tibs.2024.01.010

Prehn, K., Kazzer, P., Lischke, A., Heinrichs, M., Herpertz, S. C., & Domes, G. (2013). Effects of intranasal oxytocin on pupil dilation indicate increased salience of socioaffective stimuli. Psychophysiology, 50(6), 528–537. https://doi.org/10.1111/psyp.12042

Pruessner, J. C., Kirschbaum, C., Meinlschmid, G., & Hellhammer, D. H. (2003). Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. Psychoneuroendocrinology, 28(7), 916–931. https://doi.org/10.1016/S0306-4530(02)00108-7

Putnam, P. T., & Chang, S. W. (2022a). Interplay between the oxytocin and opioid systems in regulating social behaviour. Philosophical Transactions of the Royal Society B, 377(1858), 20210050.

Putnam, P. T., & Chang, S. W. C. (2022b). Interplay between the oxytocin and opioid systems in regulating social behaviour. Philosophical Transactions of the Royal Society B: Biological Sciences, 377(1858), 20210050. https://doi.org/10.1098/rstb.2021.0050

Putnam, P. T., & Chang, S. W. C. (2022c). Oxytocin does not stand alone. Philosophical Transactions of the Royal Society B: Biological Sciences, 377(1858), 20210047. https://doi.org/10.1098/rstb.2021.0047

Quintana, D. S. (2018). Revisiting non-significant effects of intranasal oxytocin using equivalence testing. Psychoneuroendocrinology, 87, 127–130. https://doi.org/10.1016/j.psyneuen.2017.10.010

Quintana, D. S., Lischke, A., Grace, S., Scheele, D., Ma, Y., & Becker, B. (2021). Advances in the field of intranasal oxytocin research: Lessons learned and future directions for clinical research. Molecular Psychiatry, 26(1), 80–91. https://doi.org/10.1038/s41380-020-00864-7

Quintana, D. S., Westlye, L. T., Alnæs, D., Rustan, Ø. G., Kaufmann, T., Smerud, K. T., Mahmoud, R. A., Djupesland, P. G., & Andreassen, O. A. (2016). Low dose intranasal oxytocin delivered with Breath Powered device dampens amygdala response to emotional stimuli: A peripheral effect-controlled within-subjects randomized dose-response fMRI trial. Psychoneuroendocrinology, 69, 180–188. https://doi.org/10.1016/j.psyneuen.2016.04.010

Quintana, D. S., Westlye, L. T., Rustan, Ø. G., Tesli, N., Poppy, C. L., Smevik, H., Tesli, M., Røine, M., Mahmoud, R. A., Smerud, K. T., Djupesland, P. G., & Andreassen, O. A. (2015). Low-dose oxytocin delivered intranasally with Breath Powered device affects social-cognitive behavior: A randomized four-way crossover trial with nasal cavity dimension assessment. Translational Psychiatry, 5(7), e602–e602. https://doi.org/10.1038/tp.2015.93

R Core Team. (2022). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing). https://www.R-project.org/

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: The effectiveness of reinforcement and non-reinforcement. Classical Conditioning: Current Research and Theory.

Rilling, J. K., DeMarco, A. C., Hackett, P. D., Thompson, R., Ditzen, B., Patel, R., & Pagnoni, G. (2012). Effects of intranasal oxytocin and vasopressin on cooperative behavior and associated brain activity in men. Psychoneuroendocrinology, 37(4), 447–461. https://doi.org/10.1016/j.psyneuen.2011.07.013

Rimmele, U., Hediger, K., Heinrichs, M., & Klaver, P. (2009). Oxytocin Makes a Face in Memory Familiar. Journal of Neuroscience, 29(1), 38–42. https://doi.org/10.1523/JNEUROSCI.4260-08.2009

Risko, E. F., Richardson, D. C., & Kingstone, A. (2016). Breaking the fourth wall of cognitive science: Real-world social attention and the dual function of gaze. Current Directions in Psychological Science, 25(1), 70–74.

Rogers, S. L., Speelman, C. P., Guidetti, O., & Longmuir, M. (2018). Using dual eye tracking to uncover personal gaze patterns during social interaction. Scientific Reports, 8(1), 4271.

Ross, D. L., Klykylo, W. M., & Hitzemann, R. (1987). Reduction of elevated CSF beta-endorphin by fenfluramine in infantile autism. Pediatric Neurology, 3(2), 83–86. https://doi.org/10.1016/0887-8994(87)90032-4

Roy, A., Roy, M., Deb, S., Unwin, G., & Roy, A. (2015). Are opioid antagonists effective in attenuating the core symptoms of autism spectrum conditions in children: A systematic review. Journal of Intellectual Disability Research, 59(4), 293–306. https://doi.org/10.1111/jir.12122

Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., & Baron-Cohen, S. (2015). Measuring autistic traits in the general population: A systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. Molecular Autism, 6, 2. https://doi.org/10.1186/2040-2392-6-2

Sandi, C., & Haller, J. (2015). Stress and the social brain: Behavioural effects and neurobiological mechanisms. Nature Reviews Neuroscience, 16(5), 290–304. https://doi.org/10.1038/nrn3918

Scheele, D., Kendrick, K. M., Khouri, C., Kretzer, E., Schläpfer, T. E., Stoffel-Wagner, B., Güntürkün, O., Maier, W., & Hurlemann, R. (2014). An oxytocin-induced facilitation of neural and emotional responses to social touch correlates inversely with autism traits. Neuropsychopharmacology, 39(9), 2078–2085.

Scheele, D., Wille, A., Kendrick, K. M., Stoffel-Wagner, B., Becker, B., Güntürkün, O., Maier, W., & Hurlemann, R. (2013). Oxytocin enhances brain reward system responses in men viewing the face of their female partner. Proceedings of the National Academy of Sciences, 110(50), 20308–20313. https://doi.org/10.1073/pnas.1314190110

Schilbach, L. (2015). Eye to eye, face to face and brain to brain: Novel approaches to study the behavioral dynamics and neural mechanisms of social interactions. Current Opinion in Behavioral Sciences, 3, 130–135. https://doi.org/10.1016/j.cobeha.2015.03.006

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. Behavioral and Brain Sciences, 36(4), 393–414. https://doi.org/10.1017/S0140525X12000660

Senju, A., & Johnson, M. H. (2009). Atypical eye contact in autism: Models, mechanisms and development. Neuroscience & Biobehavioral Reviews, 33(8), 1204–1214. https://doi.org/10.1016/j.neubiorev.2009.06.001

Shamay-Tsoory, S. G., & Abu-Akel, A. (2016). The social salience hypothesis of oxytocin. Biological Psychiatry, 79(3), 194–202.

Shibuki, K., Leng, G., & Way, S. (1988). Effects of naloxone and of intraperitoneal hypertonic saline upon oxytocin release and upon supraoptic neuronal activity. Neuroscience Letters, 88(1), 75–80. https://doi.org/10.1016/0304-3940(88)90318-7

Spielberger, C. D., Gonzalez-Reigosa, F., Martinez-Urrutia, A., Natalicio, L. F., & Natalicio, D. S. (1971). The state-trait anxiety inventory. Revista Interamericana de Psicologia/Interamerican Journal of Psychology, 5(3 & 4). http://journal.sipsych.org/index.php/IJP/article/view/620

Striepens, N., Scheele, D., Kendrick, K. M., Becker, B., Schäfer, L., Schwalba, K., Reul, J., Maier, W., & Hurlemann, R. (2012). Oxytocin facilitates protective responses to aversive social stimuli in males. Proceedings of the National Academy of Sciences, 109(44), 18144–18149. https://doi.org/10.1073/pnas.1208852109

Stuart, N., Whitehouse, A., Palermo, R., Bothe, E., & Badcock, N. (2022). Eye gaze in autism spectrum disorder: A review of neural evidence for the eye avoidance hypothesis. Journal of Autism and Developmental Disorders, 1–22.

Taylor, J. L., & Corbett, B. A. (2014). A review of rhythm and responsiveness of cortisol in individuals with autism spectrum disorders. Psychoneuroendocrinology, 49, 207–228. https://doi.org/10.1016/j.psyneuen.2014.07.015

Tordjman, S., Anderson, G. M., Botbol, M., Brailly-Tabard, S., Perez-Diaz, F., Graignic, R., Carlier, M., Schmit, G., Rolland, A.-C., Bonnot, O., Trabado, S., Roubertoux, P., & Bronsard, G. (2009). Pain Reactivity and Plasma β-Endorphin in Children and Adolescents with Autistic Disorder. PLOS ONE, 4(8), e5289. https://doi.org/10.1371/journal.pone.0005289

Trezza, V., Baarendse, P. J. J., & Vanderschuren, L. J. M. J. (2010). The pleasures of play: Pharmacological insights into social reward mechanisms. Trends in Pharmacological Sciences, 31(10), 463–469. https://doi.org/10.1016/j.tips.2010.06.008

Troisi, A., Frazzetto, G., Carola, V., Di Lorenzo, G., Coviello, M., D'Amato, F. R., Moles, A., Siracusano, A., & Gross, C. (2011). Social hedonic capacity is associated with the A118G polymorphism of the mu-opioid receptor gene (OPRM1) in adult healthy volunteers and psychiatric patients. Social Neuroscience, 6(1), 88–97. https://doi.org/10.1080/17470919.2010.482786

Trøstheim, M., Eikemo, M., Haaker, J., Frost, J. J., & Leknes, S. (2022). Opioid Antagonism in Humans: A Primer on Optimal Dose and Timing for Central Mu-Opioid Receptor Blockade. 1–27.

Vabalas, A., & Freeth, M. (2016). Brief report: Patterns of eye movements in face to face conversation are associated with autistic traits: Evidence from a student sample. Journal of Autism and Developmental Disorders, 46, 305–314.

Valentino, R. J., & Volkow, N. D. (2018). Untangling the complexity of opioid receptor function. Neuropsychopharmacology, 43(13), 2514–2520.

Valtakari, N. V., Hooge, I. T. C., Viktorsson, C., Nyström, P., Falck-Ytter, T., & Hessels, R. S. (2021). Eye tracking in human interaction: Possibilities and limitations. Behavior Research Methods, 53(4), 1592–1608. https://doi.org/10.3758/s13428-020-01517-x

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: Predictive coding in autism. Psychological Review, 121(4), 649–675. https://doi.org/10.1037/a0037665

Vanderschuren, L. J. M. J., Niesink, R. J. M., Spruijt, B. M., & Van Ree, J. M. (1995). μ- and κ-opioid receptor-meiated opioid effects on social play in juvenile rats.

European Journal of Pharmacology, 276(3), 257–266. https://doi.org/10.1016/0014-2999(95)00040-R

Vandesande, F., & Dierickx, K. (1975). Identification of the vasopressin producing and of the oxytocin producing neurons in the hypothalamic magnocellular neurosecretory system of the rat. Cell and Tissue Research, 164(2), 153–162.

Vatheuer, C. C., Vehlen, A., Von Dawans, B., & Domes, G. (2021). Gaze behavior is associated with the cortisol response to acute psychosocial stress in the virtual TSST. Journal of Neural Transmission, 128(9), 1269–1278. https://doi.org/10.1007/s00702-021-02344-w

Vehlen, A., Spenthof, I., Tönsing, D., Heinrichs, M., & Domes, G. (2021). Evaluation of an eye tracking setup for studying visual attention in face-to-face conversations. Scientific Reports, 11(1), 2661. https://doi.org/10.1038/s41598-021-81987-x

Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, 1, I–I. https://doi.org/10.1109/CVPR.2001.990517

Vishne, G., Jacoby, N., Malinovitch, T., Epstein, T., Frenkel, O., & Ahissar, M. (2021). Slow update of internal representations impedes synchronization in autism. Nature Communications, 12(1), 5439. https://doi.org/10.1038/s41467-021-25740-y

von der Lühe, T., Manera, V., Barisic, I., Becchio, C., Vogeley, K., & Schilbach, L. (2016). Interpersonal predictive coding, not action perception, is impaired in autism. Philosophical Transactions of the Royal Society B: Biological Sciences, 371(1693), 20150373. https://doi.org/10.1098/rstb.2015.0373

Vossel, S., Bauer, M., Mathys, C., Adams, R. A., Dolan, R. J., Stephan, K. E., & Friston, K. J. (2014). Cholinergic Stimulation Enhances Bayesian Belief Updating in the Deployment of Spatial Attention. The Journal of Neuroscience, 34(47), 15735–15742. https://doi.org/10.1523/JNEUROSCI.0091-14.2014

Vuong, C., Van Uum, S. H. M., O'Dell, L. E., Lutfy, K., & Friedman, T. C. (2010). The effects of opioids and opioid analogs on animal and human endocrine systems. Endocrine Reviews, 31(1), 98–132. https://doi.org/10.1210/er.2009-0009

Waldhoer, M., Bartlett, S. E., & Whistler, J. L. (2004). Opioid Receptors. Annual Review of Biochemistry, 73(Volume 73, 2004), 953–990. https://doi.org/10.1146/annurev.biochem.73.011303.073940

Wardle, M. C., Bershad, A. K., & de Wit, H. (2016). Naltrexone alters the processing of social and emotional stimuli in healthy adults. Social Neuroscience, 11(6), 579–591. https://doi.org/10.1080/17470919.2015.1136355

Watanabe, T., Kuroda, M., Kuwabara, H., Aoki, Y., Iwashiro, N., Tatsunobu, N., Takao, H., Nippashi, Y., Kawakubo, Y., & Kunimatsu, A. (2015). Clinical and neural effects of six-week administration of oxytocin on core symptoms of autism. Brain, 138(11), 3400–3412.

Way, B. M., Taylor, S. E., & Eisenberger, N. I. (2009). Variation in the μ-opioid receptor gene (OPRM1) is associated with dispositional and neural sensitivity to social rejection. Proceedings of the National Academy of Sciences, 106(35), 15079–15084. https://doi.org/10.1073/pnas.0812612106

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. https://ggplot2.tidyverse.org

Williams, J. R., Insel, T. R., Harbaugh, C. R., & Carter, C. S. (1994). Oxytocin Administered Centrally Facilitates Formation of a Partner Preference in Female Prairie Voles (Microtus ochrogaster). Journal of Neuroendocrinology, 6(3), 247–250. https://doi.org/10.1111/j.1365-2826.1994.tb00579.x

Williams, P. G., Allard, A., Sears, L., Dalrymple, N., & Bloom, A. S. (2001). Brief report: Case reports on naltrexone use in children with autism: controlled observations regarding benefits and practical issues of medication management. Journal of Autism and Developmental Disorders, 31(1), 103.

Windle, R. J., Shanks, N., Lightman, S. L., & Ingram, C. D. (1997). Central Oxytocin Administration Reduces Stress-Induced Corticosterone Release and Anxiety Behavior in Rats*. Endocrinology, 138(7), 2829–2834. https://doi.org/10.1210/endo.138.7.5255

Yatawara, C. J., Einfeld, S. L., Hickie, I. B., Davenport, T. A., & Guastella, A. J. (2016). The effect of oxytocin nasal spray on social interaction deficits observed in young children with autism: A randomized clinical crossover trial. Molecular Psychiatry, 21(9), 1225–1231. https://doi.org/10.1038/mp.2015.162

Yon, D., & Frith, C. D. (2021). Precision and the Bayesian brain. Current Biology, 31(17), R1026–R1032. https://doi.org/10.1016/j.cub.2021.07.044

Young, L. J., & Barrett, C. E. (2015). Can oxytocin treat autism? Science, 347(6224), 825–826. https://doi.org/10.1126/science.aaa8120

Young, L. J., & Wang, Z. (2004). The neurobiology of pair bonding. Nature Neuroscience, 7(10), 1048–1054. https://doi.org/10.1038/nn1327

Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. IEEE Signal Processing Letters, 23(10), 1499–1503.

ANNEX

Abstracts

English abstract

Autism spectrum disorder (ASD) is characterized by profound difficulties in social interaction and communication. Despite substantial efforts, underlying neurobiological mechanisms remain poorly understood and clinical interventions to alleviate social symptoms are limited. One prominent line of research has focused on the role of the endogenous opioid system in regulating social behavior and bonding. According to the Brain Opioid Theory of Social Attachment (BOTSA), excessive opioid activity may suppress social motivation in ASD, a mechanism that could be pharmacologically targeted via opioid antagonists such as naltrexone. However, larger clinical trials with naltrexone have produced inconsistent findings. A more recent perspective emphasizes the interaction between opioids and the neuropeptide oxytocin, a hypothalamic nonapeptide that acts both peripherally as a hormone and centrally as a neuromodulator, and has similarly been implicated in the regulation of social behavior. The major goal of this thesis is to investigate how interactions between opioid and oxytocin modulate social motivation and attention in ASD. The work builds on theoretical frameworks from social attention, reinforcement learning, and predictive processing, and combines mobile eye tracking, pharmacological intervention, and computational modeling. Across three empirical studies, I addressed important limitations in the existing literature: the over-reliance on screen-based paradigms in social attention research, ambiguity regarding the domain-specificity of social learning differences in ASD and the limited investigation of neurochemical interactions. In the first study, eye-tracking data were collected during naturalistic, face-to-face conversations using mobile eye tracking and computer vision–based face detection. Autistic traits were negatively associated with eye-directed gaze, supporting the ecological validity of this approach. The second study employed a reinforcement learning paradigm using point-light displays (PLDs) of social and non-social stimuli. Computational modeling was based on an extended Rescorla-Wagner framework including a reward sensitivity parameter. Higher autistic traits were associated with reduced reward sensitivity across both social and non-social conditions. This finding suggests a general learning alteration rather than a domain-specific social deficit. The third study explored the interactive effects of naltrexone and oxytocin on social gaze in individuals with and without ASD. Although overall drug effects were subtle, group differences in eye-directed gaze were attenuated under pharmacological modulation. Furthermore, cortisol responses to naltrexone were differentially associated with gaze behavior across groups: higher cortisol was linked to increased social gaze in ASD, but decreased gaze in neurotypicals. These opposing patterns suggest that opioidergic modulation of social attention may be mediated, at least in part, by interactions with the hypothalamic-pituitary-adrenal (HPA) axis. Taken together, this work advances our understanding of the neurobiological mechanisms underlying social motivation in ASD and contributes important methodological innovations for future research.

German Abstract

Autismus-Spektrum-Störungen (ASS) sind durch tiefgreifende Beeinträchtigungen in der sozialen Interaktion gekennzeichnet. Trotz intensiver Forschungsbemühungen sind die zugrunde liegenden neurobiologischen Mechanismen bislang nur unzureichend verstanden und die therapeutischen Möglichkeiten zur Unterstützung bei sozialen Beeinträchtigungen sind begrenzt. Ein zentraler Forschungsansatz befasst sich mit der Rolle des endogenen Opioidsystems in der Regulation von Sozialverhalten und Bindung. Das Opioidsystem ist nicht nur für Schmerzregulation und Belohnungsverarbeitung entscheidend, sondern spielt auch eine wichtige Rolle bei der sozialen Motivation. Basierend auf Tiermodellen und der Brain Opioid Theory of Social Attachment (BOTSA) wurde bisher angenommen, dass eine übermäßige Opioidaktivität bei ASS zur Reduktion sozialer Motivation beitragen könnte, ein Mechanismus, der durch Opioidantagonisten wie Naltrexon pharmakologisch modulierbar wäre. Größere klinische Studien mit Naltrexon lieferten bislang jedoch teilweise widersprüchliche Ergebnisse. Ein neuerer Ansatz betont die Wechselwirkung von Opioiden mit dem Neuropeptid Oxytocin. Oxytocin ist ein im Hypothalamus gebildetes Nonapeptid, das peripher als Hormon und zentral als Neuromodulator wirkt. Es wird ebenfalls mit der Regulation von Sozialverhalten und Bindung in Verbindung gebracht. Ziel dieser Arbeit ist daher, zu untersuchen, wie die Interaktion zwischen dem Opioid- und dem Oxytocin-System soziale Aufmerksamkeit und Motivation bei ASS beeinflusst. Theoretisch stützt sich die Arbeit auf Konzepte der sozialen Aufmerksamkeit, des Verstärkungslernens und des *predictive processing* Ansatzes. Methodisch wurden Eye-Tracking sowie Computational Modeling verwendet. Durch drei empirische Studien wurden wesentliche Probleme in der bestehenden Forschung aufgegriffen: Die vorherrschende Nutzung bildschirmbasierter Experimente in der Erforschung sozialer Aufmerksamkeit, die Frage nach der Spezifität von Lernunterschieden bei ASS sowie die bislang wenig erforschte Interaktion neurobiologischer Systeme. In der ersten Studie wurde das Blickverhalten von Teilnehmern während natürlicher Gespräche mittels mobilem Eye-Tracking und automatischer Gesichtserkennung mit Computer-Vision-Algorithmen erfasst. Im Vergleich zu herkömmlichen bildschirmbasierten Laborexperimenten erlaubt das naturalistische Design eine ökologisch validere Erfassung von sozialer Aufmerksamkeit. Autistische Persönlichkeitsmerkmale waren hierbei negativ mit Fixationen auf die Augenregion korreliert, was die Sensitivität des experimentellen Paradigmas für Unterschiede im Blickverhalten unterstreicht. Die zweite Studie untersuchte Verstärkungslernen mit Point-Light-Displays (PLDs). Um Unterschiede in der sensorischen Verarbeitung zu minimieren, wurden soziale und nicht-soziale Belohnungsstimuli erstellt, die hinsichtlich grundlegender visueller Eigenschaften aufeinander abgestimmt waren. Computational Modeling erfolgte auf Basis von Rescorla-Wagner-Modellen, wurde jedoch um den zusätzlichen Parameter der Belohnungssensitivität (*reward sensitivity*) erweitert. Es zeigte sich, dass höhere Ausprägungen autistischer Merkmale unabhängig vom sozialem Inhalt der Stimuli mit einer reduzierten Belohnungssensitivität einhergingen. Dies ist eher mit allgemeinen Lernunterschieden vereinbar als mit einer selektiven sozialen Lernbeeinträchtigung. In der dritten Studie wurden die kombinierten Effekte von Naltrexon und Oxytocin auf das soziale Blickverhalten bei Personen mit und ohne ASS untersucht. Zwar zeigten sich keine signifikanten Haupteffekte der pharmakologischen Bedingungen, jedoch reduzierten sich Gruppenunterschiede im Blickverhalten. Zudem zeigten sich unterschiedliche Zusammenhänge zwischen Cortisolreaktionen auf Naltrexon und Blickverhalten in beiden Gruppen: Während ein höherer Cortisolanstieg mit erhöhter sozialer Aufmerksamkeit bei Personen mit ASS einherging, war derselbe Effekt bei Kontrollpersonen negativ. Die Ergebnisse legen nahe, dass Opioide und Oxytocin gemeinsam (möglicherweise über Interaktionen mit der Hypothalamus-Hypophysen-

Nebennierenachse) soziale Aufmerksamkeit und Belohnungsverarbeitung bei ASS modulieren. Die vorliegende Arbeit leistet damit einen Beitrag zum besseren Verständnis der neurobiologischen Grundlagen sozialer Motivation und Aufmerksamkeit bei ASS und bietet methodische Impulse für zukünftige Forschung.