University of Glasgow | School of Psychology

Evaluation of the High Inflexible Precision of Prediction Errors in Autism Theory Using Simple and Biological Motion Paradigms

Greta Krasimirova Todorova (M.Sc., B.Sc.)

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

October 2021

School of Psychology
University of Glasgow
62 Hillhead Street
Glasgow
G12 8QB

University of Glasgow | School of Psychology

Declaration of Originality Form

This form **must** be completed and signed and submitted with all assignments. Please complete the information below (using BLOCK CAPITALS).

Name **Greta Krasimirova Todorova**.....
Student Number **xxxxxxx** .....
Course Name **PhD in Psychology** .....
Assignment Number/Name: **Evaluation of the High Inflexible Precision of Prediction Errors in Autism Theory using Simple and Biological Motion Paradigms**

An extract from the University's Statement on Plagiarism is provided overleaf. Please read carefully THEN read and sign the declaration below.

I confirm that this assignment is my own work and that I have:	
Read and understood the guidance on plagiarism in the Student Handbook, including the University of Glasgow Statement on Plagiarism	☒
Clearly referenced, in both the text and the bibliography or references, **all sources** used in the work	☒
Fully referenced (including page numbers) and used inverted commas for **all text quoted** from books, journals, web etc. (Please check with the Department which referencing style is to be used)	☒
Provided the sources for all tables, figures, data etc. that are not my own work	☒
Not made use of the work of any other student(s) past or present without acknowledgement. This includes any of my own work, that has been previously, or concurrently, submitted for assessment, either at this or any other educational institution, including school (see overleaf at 31.2)	☒
Not sought or used the services of any professional agencies to produce this work	☒
In addition, I understand that any false claim in respect of this work will result in disciplinary action in accordance with University regulations	☒

DECLARATION:
I am aware of and understand the University's policy on plagiarism and I certify that this assignment is my own work, except where indicated by referencing, and that I have followed the good academic practices noted above

Signed: .....

Abstract

Autism Spectrum Disorder is a lifelong neurodevelopmental condition associated with large lifetime costs to services, the individuals and their families. To develop appropriate support, a unified account of the condition needs to be found to provide a framework for research to build on. This thesis focuses on testing one of the recently proposed theories – the theory of High Inflexible Precision of Prediction Errors in Autism (HIPPEA), which interprets autism through a predictive coding perspective. The predictive coding framework argues that through experience, the brain forms predictions about the incoming sensory information, which it then compares with the actual input. Mismatches produce prediction errors which are weighed in comparison to the prediction, and if enough weight is assigned to the precision of the prediction error, a change in the prediction or the action is enacted. HIPPEA poses that autism arises from a difference in the tuning of this general neurocognitive mechanism whereby attention leads to the invariably high precision setting of prediction errors. This, in turn, leads to the creation of narrow prediction models that are based on infrequent contingencies and noise. This PhD aims to contribute research results and paradigm designs that investigate precision weight setting of prediction errors in autism.

This thesis presents three behavioural and one neuroimaging experiments, and one meta-analysis. Each experiment modulates attention and expectation under different experimental paradigms allowing for the investigation of these two factors in multiple contexts. Chapter 2 makes use of an established apparent motion paradigm. In this chapter, endogenous attention is controlled allowing the investigation of the differences in prediction establishment and prediction error processing in neurotypical and autistic individuals. Moving forward, to establish the viability of using biological motion stimuli as an effective way to measure differences between autistic and non-autistic individuals, Chapter 3 presents a large-scale meta-analysis of behavioural, eye-tracking, EEG and fMRI studies investigating biological motion perception and interpretation in autism. Chapter 4 presents two studies that look at the effects of autistic traits in a task that orthogonally modulates attention and expectation by explicitly instructing participants about the statistical regularity of events and by providing implicit

cuing using a human point-light kicker or a coherent dot-motion display. Finally, Chapter 5 presents a proof-of-concept study, which examines the feasibility of a modification in a recently developed EEG paradigm of hierarchical frequency tagging of bottom-up and top-down signals using dynamic human biological motion. This paradigm allows the investigation of the representation of low- and high-level components of the human point-light display, along with their integration in the brain while modulating attention and expectation through task instruction.

The results from this thesis indicate that like neurotypical participants, autistic individuals can create and benefit from the development of predictions either through illusory motion or through the explicit establishment of expectations. In line with HIPPEA, this indicates that it is not the establishment of predictions that is the cause of the traits observed in autism. Moreover, what we see is that unpredictable events are treated differently, suggesting disproportionate amplification of unpredictable events, as suggested by HIPPEA. However, we do not see support for the 'inflexible' part of the HIPPEA theory. Instead, this thesis concludes that prediction errors show some special treatment in autism, but that is context-dependent. For research to move forward, it is paramount that attention is a controlled factor, and that context-dependent precision weight setting of prediction errors is incorporated in a reviewed version of the HIPPEA theory.

Table of Contents

List of Tables

List of Figures

Previous Publications

Chapters 2, 3, and 4 are presented in manuscript form and are either published, under review or in preparation.

Chapter 2

Todorova, G.K, Pollick., F. E & Muckli, L. (under review). Special treatment of prediction errors in Autism Spectrum Disorder. Neuropsychologia.

Chapter 3

Todorova, G.K, Hatton, R.E.M & Pollick., F. E. (2019) Biological motion perception in Autism Spectrum Disorder: A meta-analysis. *Molecular Autism,10*(1), 49. DOI: 10.1186/s13229-019-0299-8

and

Todorova, G. K., Hatton, R. E. M., & Pollick, F. E. (2021). Correction to: Biological motion perception in autism spectrum disorder: a meta-analysis. *Molecular Autism 2021 12:1*, *12*(1), 1–3. https://doi.org/10.1186/S13229-021-00462-6

Chapter 4

Todorova, G.K. & Pollick, F.E. (in preparation). Effects of Attention and Predictability are not Affected in Individuals with High Autistic Traits: Two Dynamic Online Experiments. Journal of Vision.

Acknowledgements

This thesis would not have been possible without the support of my supervisors. I would like to thank my Primary supervisor Frank Pollick, for putting me in contact with specialists in different fields, making my PhD process smoother; for the weekly meetings, which were never daunting, and always felt like a supportive and friendly environment; for believing in me and trusting my reasoning process; and most importantly, for the pastoral support when things were not going great, when my back gave out for a year, and I couldn't work to the same capacity, and during a 1.5-year pandemic. The mark of a great supervisor is when the student knows that they can count on them at any point. I would also like to thank my Secondary supervisor Lars Muckli for challenging my arguments and reasoning and his helpful comments on my written work.

I would also like to thank the researchers who shared their time and knowledge with me because those are the academics that made me like academia: Dr Christoph Scheepers and Dr Dale Barr for sharing their immense stats knowledge; Dr Christopher Benwell, Dr Noam Gordon, Asst Prof Jeroen Van Boxtel and Prof Jacob Howhy for their invaluable help during the set-up of the EEG project in Chapter 5, which was so devastatingly interrupted by the COVID-19 pandemic; Prof Ian Thornton for his advice during the development of the paradigm in Chapter 4. And of course, the PhDs at the School of Psychology who shared life and academic wisdom – you were an excellent support system.

This PhD would not have been possible without the participants who shared their time by participating in the research projects and the funders (SGSSS/ESRC and SINAPSE), who believed in my projects.

And finally, I would like to thank my partner and my family. I could not have done it without your emotional support throughout the whole process.

Abbreviations

Abbreviation	Meaning
(med)SFG	(medial)Superior Frontal Gyrus
(p)STS	(posterior) Superior Temporal Sulcus
ACG	Anterior Cingulate Gyrus
ADOS	Autism Diagnostic and Observation Schedule
AG	Angular Gyrus
ALE	Activation Likelihood Estimation
AQ	Autism-Spectrum Quotient
AR	Action Recognition
ASD	Autism Spectrum Disorder
BA	Broadman Area
BM	Biological Motion
CI	Confidence Interval
CM	Coherent Motion
COH	Coherent Dot Motion
D	Biological Motion Detection
DSM	Diagnostic and Statistical Manual of Mental Disorders
EEG	Electroencephalogram
EHI	Edinburg Handedness Inventory
ER	Emotion Recognition
FFG	Fusiform Gyrus
FFT	Fast Fourier Transform
FLD	Full-Light Display
fMRI	Functional Magnetic Resonance Imaging
FrACT	Freiburg Vision Test
FSIQ	Full-Scale Intelligence Quotient
HAT/LAT	High/Low Autistic Traits
HFT	Hierarchical Frequency Tagging
HIPPEA	High Inflexible Precision of Prediction Errors in Autism
IFG	Inferior Frontal Gyrus
IM	Intermodulation
IOG	Inferior Occipital Gyrus
IOR	Inhibition of Return
IPL	Inferior Parietal Lobule
ITG	Inferior Temporal Gyrus
L	Left
MAD	Median Absolute Deviance
MCG	Middle Cingulate Gyrus
MFG	Middle Frontal Gyrus
MMN	Mismatch Negativity
MNN	Mirror Neuron Network
MOG	Middle Occipital Gyrus
MSPC	Multi Spectra Phase Coherence
MT	Middle Temporal Area
MTG	Middle Temporal Gyrus
N	Sample Size
NT	Neurotypically Developing
NVIQ	Non-Verbal Intelligence Quotient
PABAK	Prevalence-Adjusted and Bias-Adjusted Kappa
PLD	Point-Light Display

Abbreviation	Meaning
PT	Predictable Target
PTS	Posterior Temporal Sulcus
PV	Passive Viewing
R	Right
RT	Reaction Time
SCR	Scrambled Biological Motion
SNR	Signal-To-Noise Ratio
SOA	Stimulus Onset Asynchrony
SQA	Standard Quality Assessment Score
SSVEP	Steady State Visually Evoked Potential
STG	Superior Temporal Gyrus
SWIFT	Semantic Wavelet-Induced Frequency-Tagging
TMS	Transcranial Magnetic Stimulation
TPJ	Temporal-Parietal Junction
UT	Unpredictable Target
vlPC	Ventrolateral Prefrontal Cortex
vmPC	Ventromedial Prefrontal Cortex
WASI	Wechsler Abbreviated Scale of Intelligence
WB	Whole Brain Analysis
WoE	Weight of Evidence

Chapter 1 Autism Spectrum Disorders and the predictive processing framework

Autism Spectrum Disorder (ASD) is a lifelong neurodevelopmental condition characterised by restrictive and repetitive behaviours and social and communicative difficulties, which can but do not have to co-occur with language difficulty (American Psychiatric Association, 2013). Additionally, one of the new additions to autism characterisation in the final updates of both the International Classification of Diseases -11 (World Health Organisation, 2020) and Diagnostic and Statistical Manual of Mental Disorders – 5 (DSM 5) (American Psychiatric Association, 2013) is the presence of hyper- and/or hypo-sensitivities or reactivities to sensory information. Prevalence statistics vary between years, countries and methodology used to obtain the statistics, which obscures the condition's true prevalence (Fombonne, 2018). Prevalence measures range from 1.68% in 8-year-olds in the US (Baio et al., 2018), to ~1% in China (Sun et al., 2019), 1.1% in England (Brugha et al., 2012), between 0.4 – 1.9% amongst countries in the European Union (ASDEU, 2018) and 2.64% in South Korea (Kim et al., 2011). Overall, it appears that autism prevalence tends to be slightly above 1% across countries but with a large standard deviation. As the condition is a lifelong condition, the lifetime costs to society in the UK range between £1.5 million and £2.4 million, with 56% of costs accounted for by services, 42% by loss of employment and 2% in caregiver costs (Rogge & Janssen, 2019). In the US, lifetime social costs (medical and non-medical costs) of the condition are approximated to be around $3.6 million in 2019, with the total cost between 1990-2019 approximated to ~$7 trillion and expected to reach between $11.5 trillion if the prevalence stays the same and $14.9 trillion if the prevalence increases in a similar rate to previous decades (Cakir et al., 2020). However, it is argued that some of these costs could be reduced if risk factors and better support measures are found (Cakir et al., 2020). To be able to achieve this, a reasonable and unified account of the condition is needed to provide a framework on which research can build on.

This thesis focuses on testing one of the recently proposed theories for understanding autism, which interprets autism through a predictive coding perspective - the theory of High Inflexible Precision of Prediction Errors in

Autism (HIPPEA, Van de Cruys et al., 2014). It poses that autism characteristics arise from a difference in the tuning of a general neurocognitive mechanism, rather than a specific dysfunction. The proposed theory puts autism at one end of the spectrum of cognition where life experiences and genetic contributions allow for a neurocognitive mechanism that functions efficiently in neurotypicals to lead to the presentation of autism. This thesis contributes to the available and accumulating evidence investigating this theory through the use of psychophysical paradigms along with new conceptual paradigms with clearly defined outcomes.

This chapter will first briefly introduce autism as a condition and then it will focus on the applicability of the predictive coding framework to autism. It will introduce predictive coding as a general framework and review recent evidence for and against two of the main competing theories, justifying the selection of one of them as the topic under investigation.

Throughout this thesis, individuals diagnosed with autism spectrum disorder will be referred to with diagnosis-first (i.e., autistic individual) and person-first (i.e., individual on the autism spectrum) labelling. As there is no existing consensus within the autism community about the use of language (Bury et al., 2020; Kenny et al., 2016) the terminology will be used interchangeably. Additionally, when discussing participants from the general population – like control participants, they will be referred to as neurotypical, rather than typical, healthy, or normal. This aims to minimise the stigma often associated with autism as unhealthy, abnormal, or atypical. Instead, it emphasises the neurodiversity of the two populations and the neurodevelopmental nature of autism.

1.1 What is autism?

Autism is a lifelong condition characterised by large heterogeneity in phenotypical as well as in genetic expression. Individuals on the spectrum can vary from needing little to no support, to requiring substantial support due to severe impairment in functioning (Masi et al., 2017). Substantial time has passed since the first description of autism by Kanner in 1943. Moreover, since autism was introduced as a separate diagnostic category three different updates of the

Diagnostic and Statistical Manual of Mental Disorders (DSM) have come and gone. However, the diagnosis of autism is still based on behavioural characteristics (American Psychiatric Association, 2013; World Health Organisation, 2020). This is mainly due to the lack of consensus in findings about the underlying cause of autism.

ASD is commonly described as a genetic condition and it is estimated that anywhere between 400 and 1000 genes can lead to its development along with a wide range of de novo copy-number variants (a genetic mutation or variation, that is not inherited, and occurs for the first time in an individual) (Masi et al., 2017). However, a large amount of the heterogeneity comes after birth during the developmental stages, potentially due to higher susceptibility linked to this varied genetic make-up (Landrigan, 2010; Masi et al., 2017). In this way, no two individuals are the same and there is not a homogenous autism phenotype. Many of the characteristics specific to autism are present to various degrees among the general population, however, it is the severity to which they occur and interact within autistic individuals that leads to impairment in their daily functioning (American Psychiatric Association, 2013). In fact, questionnaires that measure autism-like characteristics in the general population have been used as proxies to observe how varying degrees of these traits can affect task performance in the neurotypical population. Autism-like characteristics, as measured by the Autism Quotient (AQ, Baron-Cohen et al., 2001), are observed to a lesser degree within the general population, however, participants with higher scores have been shown to reliably demonstrate performance patterns similar to individuals on the autism spectrum (eg. Cribb et al., 2016; Stewart et al., 2009).

Much research has gone into attempts to characterise autism and its phenotype. In the perceptual domain, it has been shown that individuals on the autism spectrum are more likely to focus on local components as opposed to the global picture, which is evident in faster reaction times in tasks that require finding smaller patterns within a larger whole (Simmons et al., 2009). However, these findings are dependent on the task itself and the makeup of the sample (Van der Hallen et al., 2015; Van Eylen et al., 2015). Similar findings have been observed in the auditory literature, with better performance on pure-tone detection tasks but this performance tends to disappear in complex melodic tasks (O'Connor,

2012). Additionally, investigation of global motion perception abilities has indicated reduced ability to detect global motion in tasks such as coherent motion detection tasks from random-dot kinematograms (Van der Hallen et al., 2019). Attention research in autism has also suggested a difficulty in the ability to respond to rapidly presented cues, reduced divided attention and shifting attention abilities, and more spatially narrow focused attention, which is suggested to lead to longer response times, rather than complete failure in detecting targets (Ames & Fletcher-Watson, 2010). There appear to be no difficulties in alerting behaviour, but problems arise at the point of disengagement (Orekhova & Stroganova, 2014).

A significant amount of research has also tried to characterise differences in the social domain. Social difficulties are also some of the more pronounced in autism, with difficulties in maintaining joint attention, using gaze direction to engage in joint activities and figurative language comprehension, reduced attention to faces and voices (Kalandadze et al., 2018; Papagiannopoulou et al., 2014; Volkmar, 2011). Additionally, similar to the findings in the perceptual domain, findings concerning the perception and interpretation of human biological motion are also dependent on the type of task. However, there appears to be a consistent reduced ability to detect and to extract information about intentions and emotions from human movement from degraded (point-light displays) and complete stimuli (Federici et al., 2020; Todorova et al., 2019; Van der Hallen et al., 2019). Differences in these areas can, in turn, lead to difficulty in forming relationships and feelings of isolation (Muler et al., 2008; Sosnowy et al., 2019).

Restrictive and repetitive behaviours and interests are also an essential part of the autism diagnosis. They encompass a variety of behaviours such as repeatedly performed movements and sensory behaviours, areas of intense interests and insistence on sameness. Repetitive behaviours have been negatively correlated with adaptive skills (Szatmari et al., 2006). However, these behaviours appear to reduce with age (Barrett et al., 2018). On the other hand, insistence on sameness has been shown to be negatively correlated with characteristics in the communication and language domains (Szatmari et al., 2006) but it does not appear to show a change with age in adults (Barrett et al., 2018). Furthermore, the levels of these behaviours differ between the sexes, with males showing

heightened stereotyped behaviours and restricted interests, whereas females show heightened compulsive, restricted and self-injurious behaviours (Antezana et al., 2019; McFayden et al., 2020). Moreover, females tend to demonstrate more interests with social content such as animals, activism etc., which can lead to interests being missed (McFayden et al., 2020).

From a neuroscience perspective, a meta-analysis of functional magnetic resonance imaging (fMRI) studies on autism shows that there are consistent differences between autistic and neurotypical individuals (Philip et al., 2012). Some of the findings include reduced activation in visual processing areas, reduced activation in areas associated with receptive language processing, areas related to face and human form perception, decreased activation in automatic attentional networks and networks associated with cognitive control and hypo- and hyper-activation in areas related to motor tasks, as well as increased activation in areas associated with learning and planning through reward and punishment (Philip et al., 2012). Additionally, the differences in brain function appear to change with age, indicating that different networks become involved with increasing age. This finding is very important, as it is often seen in the literature to talk about individuals on the autism spectrum, rather than the different developmental stages. As developmental stages play an important role in neurotypicals, they should be considered when extrapolating to the condition as a whole (see Chapter 3; Crawley et al., 2019; Todorova et al., 2019).

It is clear that autism is a very diverse condition, with findings in multiple areas of functioning. Explaining autism and its characteristics has hence been a large area of research with numerous cognitive theoretical accounts of autism gaining and losing traction in waves throughout the years (Fletcher-Watson & Happe, 2019). A few will be briefly described below. This is not an exhaustive list of the existing theories about autism but will illustrate some of the most prominent ones.

Attempt to explain ASD have tended to focus on certain components of the autism phenotype rather than on the coherent whole. On one hand, the ‘Theory of Mind’ account as described by Baron-Cohen (2000) focuses mainly on the social component of autism. It is built around the ability to attribute separate mental states to oneself and other people. The ‘Theory of Mind’ explanation of

autism argues that shortcomings in these mentalising abilities provide an account of the social difficulties related to autism, where the reduced ability to socially interact with neurotypical individuals and to interpret social cues is due to a difficulty in developing mentalising abilities like joint attention (Baron-Cohen et al., 1992). On the other hand, the theories of Happe and Frith (2006) and Mottron et al. (2006) focus mainly on the perceptual characteristics of individuals on the autism spectrum - lower susceptibility to illusions, better performance on tasks requiring a focus on details over a focus on the global context, savant abilities. Whereas the former theory argues for a local bias with the cost to global processing, the latter argues for enhanced local processing with preserved global. Unfortunately, which one of these arguments best represents autistic individuals best is not determined due to many conflicting findings (Simmons et al., 2009).

These theories are not mutually exclusive and although they have had a great influence over the way the field develops, they are often unable to explain the characteristics that are not central to the theory (Fletcher-Watson & Happe, 2019). Thus, there is a need for a unifying concept that encompasses all components of autism, rather than focusing only on one or a few of the condition's characteristics. There have been some attempts to achieve this by explaining the condition with either too high (Simmons et al., 2007, 2009) or too low neural noise (Davis & Plaisted-Grant, 2015), with arguments that the low endogenous noise theory explains the high global noise in the autistic neural system (Simmons & Milne, 2015). However, these theories tend not to provide complete explanations of the social aspects of the condition.

1.2 New theoretical avenues for autism research

For the past several decades, our understanding of the way people interact with the world and the way the brain encodes and perceives it has been driven by the understanding that the brain is not just a passive observer and receiver of information (Friston et al., 2017). Indeed, the brain encodes a representation of our environment and that representation is consistently tested against the reality of sensory experiences (Friston et al., 2006). These models are based on Bayesian theory and the hierarchical and recurrent nature of the brain architecture and present the brain as a predictive system. The predictive

processing framework, which is based around this premise, explains human behaviour and its variations between individuals as differences in these general neurocognitive mechanisms. In this sense, the predictive processing framework could provide a unifying framework for numerous phenomena and neuropsychological conditions. One of these is autism.

The rest of this chapter will be structured in the following way. Section 1.2.1 will provide an overview of predictive processing in general. It is important to note that there are different models and propositions that fall under the 'predictive processing' framework like active inferencing as proposed by Friston et al. (2017) and the Opposing Processes as proposed by Press et al. (2020). In the section, I aim to present a broad overview of what is discussed as predictive coding, and it should be kept in mind that there are some differences in the specifics between accounts. In Section 1.2.2 and 1.2.3, the application of this framework to autism will be discussed by describing the two competing autism theories based on the predictive processing framework. Sections 1.2.4 and 1.2.5 will discuss research that tests the main concepts of the two competing theories. The final section 1.3 will bring all the information together, charting the purpose of this thesis and the way the rest of the chapters will progress.

1.2.1 The predictive processing framework

The cornerstone of the predictive processing framework is the idea that the brain attempts to maintain homeostasis and minimise entropy by making predictions about the incoming information, as opposed to other learning theories like reinforcement learning, where the proposed mechanism is the brain's attempt at maintaining a high-reward state (Friston, 2009; Sajid et al., 2019).

When a prediction error is encountered, the system can either update the predictive model, rely on the prediction or act out on the environment to change it (Friston et al., 2011). In this way, the biological system minimises the entropy introduced by prediction errors. In this context, Friston (2010) formulates entropy as a measure of uncertainty or the average surprise of outcomes. At the simplest level, we can look at the predictive processing framework as explained by Friston and colleagues in one of their seminal papers

(Friston et al., 2011). In active inference, predictions are formed based on specific goals and knowledge about the current state of the environment. Actions are executed as a response to those predictions, attempting to attain those initial goals. Mismatches between the strong predictions about the state of one's body and the external information received from the body's movement are minimised, up weighting the prediction by engaging reflexes to make sure those predictions are attained (Friston, 2010; Yon et al., 2019). On the other hand, in perceptual inference, the perceptual input that we get from the environment is used to update our predictions and in turn, minimise any future prediction errors (Yon et al., 2019). Thus, the brain learns new information about the way the world functions using information from the past and predictions about the future in combination with prediction errors when those predictions turn out to be inaccurate.

Evidence for the 'entropy minimisation' aspect of the brain comes from studies that have shown that predicted events reduce the activation in lower-level brain areas through modulatory behaviour of higher-order brain areas. Minimising activation in lower-level sensory areas in such a way allows the brain to use less energy. One example of this pattern of activation is seen in apparent motion paradigms, where stimuli that appear on the illusory motion path in time with the expected motion elicit lower primary visual cortex (V1) activation in comparison with stimuli that appear out of time (Alink et al., 2010). Moreover, this specific behaviour appears to be modulated by the activation in higher-order brain areas like V5, which are associated with motion perception (Vetter et al., 2012). Similar behaviour has been observed in cueing paradigms, where a cue indicating the location of a stimulus leads to reduced V1 activation when the stimulus occurs at the cued location in comparison to when the stimulus occurs at a different location (Kok, Rahnev, et al., 2012). Moreover, expectations formed about the specific pattern of visual stimulation, have been shown to elicit anticipatory activity in V1 after the presentation of the beginning of a sequence in the absence of the following stimuli (Ekman et al., 2017). This pattern of activation was not observed when the ending of the expected sequence was presented. This indicates that the brain creates expectations not only about locations but also about patterns that are related in space and time. Using more complex stimuli, Muckli et al. (2015) were able to show that in a

paradigm where part of a real-life stimulus was absent – a blank quadrant from an image - brain activity can be used to decode the missing part of the image. More specifically, brain activity in the outer superficial layers of V1 was indicative of contextual feedback information, whereas observing the full image showed a peak activation in mid-layers of V1. The origin of that contextual information is suggested to come from extrastriate visual areas, along with higher-order cortical and subcortical regions (Muckli et al., 2015).

One specific example of entropy minimisation has been provided by Tod and Cornwell (2018). They used a paradigm, where participants learnt to expect a specific variability of the occurrence of a deviant stimulus within a standard stream of tones – either a very small deviation from the standard onset time (±10ms) or a large deviation from the standard onset time (±200ms). Using mismatch negativity (MMN) they showed that when moving from the more precise (10ms) to the less precise (200ms) environment there was a decrease in the N2 component in the electroencephalographic (EEG) signal, which corresponds to the oddball effect in such tasks. However, the decrease was consistent with the reduction of the activation to the standard. Thus, overall, the activation difference between the standard and the deviant had not changed. The authors attribute this behaviour to the fact that the deviant's onset is less precise – further away from the expected 500ms onset-to-onset timing. This creates a noisier environment in comparison to the learned variability and the brain should not aim to incorporate the new information into the existing model as the already existing predictive model is 'good enough' and the new prediction errors would be considered noise (Todd & Cornwell, 2018). This interpretation would lead to the logical conclusion that the brain allows for small variability within the prediction to occur but when the variability becomes too large, the brain should not aim to decrease the specificity of the model as that would create a noisier prediction. When the participants moved from the 200ms deviation to the 10ms deviation – i.e., from a less to a more precise prediction, Todd and Conrwell (2018) observed an initial increase in the N2 component, but with time, the component returned to the previous difference observed when the expectation was being established. Todd and Cornwell (2018) attribute this phenomenon to an increase of precision – the shorter variation in onset is more precise than the larger because the difference from the expected

onset-to-onset timing is smaller. The change in the expectation leads to initial prediction error detection, which is later accumulated into the new prediction. In short, the prediction was updated based on encountering an error that differed from the expected variation. In summary, Todd and Cornwell (2018) aimed to show that when moving from a less to a more precise environment, even in a passive task, the brain would respond in a way that would try to minimise uncertainty and in this way either lead to updating the prediction and learn to expect new variation or to discounting the added uncertainty as noise.

Paradigms like this provide us with a lot of insight into the way the brain deals with prediction errors in a passive environment. Repeating the experiment but changing the task to an active one, could also allow us to see how these effects will change when there is an action required in the presence of an odd sound. It is possible that in that case the 200ms oddball would elicit a larger N2 component because responding to it would coincide with the goal of the system. It would also be interesting to see what would happen if there was variability in the presentation of the stimulus – i.e., comparing activation when timing varies between 100-200ms and between 10-20ms for example would make the evidence of precision more convincing. Further, this study, unfortunately, is not able to show what happens if the old conditions return to observe whether the brain has indeed updated the old prediction, or what the authors are observing is to an extent attributable to repetition suppression. Such changes, however, would have complicated the design, reduced the power, and potentially introduced noise in the EEG signal. Thus, this paradigm would have to be considered a good proxy for the proposed framework.

Hence, learning in the predictive processing framework occurs only if enough weight is given to the incoming information and minimises the uncertainty of the environment. However, if the environment does not provide enough information for the organism to be able to establish its regularity, a prediction error would not have the same amount of precision because the statistical regularity of the environment is not strong enough to evoke a high precision prediction error. The importance of the regularity of the environment for precision modulation has been shown by Southwell and Chait (2018), where a deviant sound in terms of pitch was presented in the context of a predictive or a random pattern. In their results, the deviant sound in the predictive pattern elicited a larger response in

comparison to the deviant sound in the random pattern. In this situation, a prediction error in a predictive pattern would be more precise because the environment would be less noisy. In the random pattern, a deviant sound in terms of pitch would be surprising, however as there is a less specific prediction, there would be a larger allowance for prediction errors and would elicit a smaller response than the prediction error, which is coming from an environment with a narrow prediction. Thus, not every prediction error would be due to an inaccurate prediction as it might be caused by noisy stimuli. In these cases, we observe increased weighting of the prediction, rather than the prediction error in the form of sharpening of the expectation (Kok, Jehee, et al., 2012; Press & Yon, 2019).

Thus, the overall main components of the predictive processing framework are the prediction and the prediction error. The balance between these two components and the influence that each one has on the neurocognitive system is determined by the precision setting of each one.

1.2.1.1 The hierarchical structure of the predictive processing framework

The underlying concept of the predictive processing framework is based on predictions within hierarchical systems. The brain by nature is a hierarchical structure, and at each level, there is a generative model that forms a prediction about the expected incoming information (Sajid et al., 2019). If we take the visual system in the brain, a simple overview will tell us that the information from the eyes enters the primary visual cortex (V1), then travels up the hierarchy where more and more complicated information about the visual stimulus is decoded – i.e., the fusiform face area is suggested to deal with face perception, V5/MT with motion perception, parietal place area with houses, etc. At each level of the hierarchy, the incoming information is matched against the prediction. How a mismatch between the error and prediction is handled is dependent on the precision setting of the prediction error. If the new incoming information is not ambiguous and/or the task requires a reaction or response – i.e., the prediction error is of relevance to the task at hand, then it would be of higher precision. Thus, if the prediction error is of high precision, then an

attempt will be made to update the prediction to explain away the difference between the predicted and the encountered stimulus.

Such hierarchical structure has been revealed in a cortical coupling paradigm by den Ouden et al. (2010), where the probability of a face and a house were varied throughout the course of an experimental trial. Specifically, they varied the cue-outcome associative contingency by varying the amount to which an auditory cue predicted a face or a house image. In this scenario, a prediction error would be observing a face, when the specific auditory cue would have suggested a house. fMRI results showed that when the stimuli (face/house) were not expected i.e., surprising stimuli, higher activation was observed in the fusiform face area/parietal place area, respectively. More importantly, the putamen showed stimulus-independent activation to prediction errors. Further, dynamic causal modelling showed that the putamen was facilitating the strength in connection between the premotor cortex and the fusiform face/parietal place area. Thus, these areas represent the hierarchical system engaged in the present task. The sensory information is decoded by the fusiform face/parietal place area. A larger prediction error - i.e., the less expected – the stronger the reaction. In this paradigm, when a face was presented but a house was expected and the cue-outcome association was strong, the prediction error became more relevant, as a motor task (a button press) needed to indicate the perceived stimulus. Thus, the putamen would facilitate the propagation of the prediction error up the hierarchy to the pre-motor cortex to change the motor command that was originally primed by the auditory cue. In this way, the relation between the observation of the face, its association with the auditory cue and the premotor cortex represent a hierarchical system where prediction errors occur at every level.

1.2.1.2 The role of attention in prediction

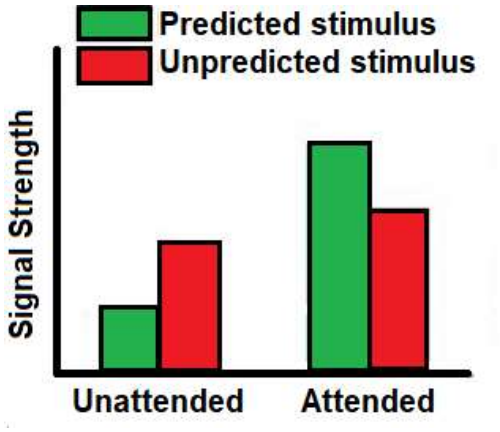
As stated earlier, there are different levels of precision setting of the prediction errors and the predictions themselves. The level of estimated precision of each of them varies depending on context and the system flexibly adjusts what information from the world needs to be learned. The parameters of these different levels of precision are called hyperparameters and the prior beliefs about them are called hyperpriors (Friston, 2008). These hyperparameters are

formed over time when the individual learns in which situation the incoming information is noisy and in which the variation from the prediction is caused by a change in the rules about the environment. One of the mechanisms through which hyperparameters work is suggested to be attention (Feldman & Friston, 2010; Friston, 2008; Parr & Friston, 2019). According to some predictive processing accounts, attention is important in uncertainty estimation and by focusing attention on a stimulus this allows an unexpected event to be estimated as a more precise prediction error that should be propagated further up the hierarchy to update the prediction, rather than explained away at earlier stages (Feldman & Friston, 2010). Attention assists in estimating the relevance of a stimulus, which leads to giving higher weight to prediction errors when they are attended by optimising synaptic gain during hierarchical inference (Feldman & Friston, 2010; Parr & Friston, 2019). Under this condition, attention and expectation need to be modulated orthogonally to allow investigating both attention and expectation separately. Several studies have attempted to achieve this, and they will be discussed below.

In a simple attention cueing paradigm, Kok, Rahnev and colleagues (2012) created an expectation in their participants about the consistency of cues appearing on the left or right side of the screen. At the same time, they also modulated attention by providing a directional cue - left/right, which indicated on which stimuli (ones appearing on the left or right side of the screen) participants had to perform an orientation judgement task. In this sense, expectation and attention were modulated separately. In the primary visual cortex (V1), predicted stimuli showed a reduced response. This type of brain activation has been observed in other prediction-oriented paradigms such as the one by Alink et al. (2010), which was discussed earlier. In Alink et al.'s (2010) paradigm, behavioural results indicating better detection of the predictable as opposed to the unpredictable stimuli showed that the increased activation for the unpredictable stimuli cannot be attributed to attention. Rather the activation pattern could be interpreted as a prediction error activation, which was not propagated further up the hierarchy as it was not attended (De-Wit et al., 2010). More importantly, Kok, Rahnev and colleagues (2012) were able to show that this pattern reverses when items are attended to. Specifically, if an item was predicted and relevant (attended) then this increased the activation in

the primary visual cortex in comparison to the attended but unpredicted stimulus. However, if the item was predicted, but it was not relevant (i.e., not attended), it led to reduced activation in comparison to the unpredicted stimuli. These results describe how attention can have a modulatory effect on the relevance or precision of prediction errors. A graphical representation of Kok, Rahnev et al.'s (2012) results is shown in **Error! Reference source not found.**.

Figure 1. Effect of attention on predicted and unpredicted stimuli (Adapted from Kok, Rahnev, et al., 2012).



Similar findings have been observed utilising a new methodology called hierarchical frequency tagging. This is a paradigm where using changes in the stimuli (flickers or cycling of images in-and-out of noise) can introduce entrainment in the neural activity of the brain, which is observed through EEG. In a set of experiments, Gordon et al. (2017; 2019) used a high-level representative stimulus – house or face, which emerged at a specific frequency from a wavelet transformation that scrambles the characteristics of the images but keeps the local luminance the same (Koenig-Robert & VanRullen, 2013). This allowed them to modulate the high-level representation in the brain. At the same time, they created a luminance flicker of the whole display at a different frequency thus, modulating the low-level information of the display. These changes in the presented stimuli then allowed them to observe the entrained brain activity to the changes in the presentation of the high-level semantic images while at the same time observing the brain activity with the change of the more low-level characteristic changes of the stimulus (luminance). Moreover, by modulating the proportion of houses and faces being presented (expectation modulation) and focusing attention towards counting the faces or the houses, Gordon, Tsuchiya et al. (2019) were able to separately modulate

expectation and attention. What they observed was that expectation modulation was associated with a decrease of the EEG signal associated with the house/face image (Gordon et al., 2017, 2019; but see Coll et al., 2020). More importantly, expectation modulated the integration of these two stimuli, as represented by changes in the intermodulation (IM) components, which are observed when waves at two frequencies are combined. The modulation of these IM components was in such direction that expectation leads to an increase in their signal-to-noise ratio (Gordon et al., 2017).

To investigate attention, Gordon, et al. (2019) calculated the extent to which the IM components' phase synchronised with the phase of the original stimulus, or with the phase of the EEG signal. The authors proposed that these two metrics represent the interactions occurring at lower or higher levels of the hierarchy, respectively. They observed that the extent to which the IM components synchronised with the stimulus (MSPCstim) was affected by expectation with an increase of synchronisation with higher expectation. No modulation of attention was observed for MSPCstim. The opposite was observed for the metric that represented the synchronisation of the IM components with the EEG signal (MSPCres). Moreover, with the increase of predictability of the occurring image (house or face), MSPCres showed an increase for attended and a decrease for unattended images. Gordon and colleagues (2019; 2019) argue that the effect of attention increases the integration of low-level and high-level information at higher levels of the hierarchy only if the stimuli are attended to. These results corroborate the findings of the effect of attention as an important modulating factor of precision in Kok, Rahnev et al. (2012). However, it must be kept in mind that the use of the MSPCstim and MSPCres is still in its early stages. Despite the repeated success in finding similar results with different equipment (Gordon, Tsuchiya, et al., 2019), a recent attempt of replicating the result with a different set of images (Coll et al., 2020), suggests that not all of the effects might be generalisable above and beyond the specific stimulus set. Thus, more research is needed to determine the specificity of the proposed methodology.

It is important to mention, that it is not necessarily clear whether attention would increase accuracy or would just act as a gain control mechanism increasing the synaptic response irrespective of accuracy (Maunsell, 2015; Mehrpour et al., 2020). However, a recent study on rhesus monkeys attempts to

provide insight into this (Mehrpour et al., 2020). Mehrpour et al. (2020) observed a population of neurons in the brain that are sensitive to directional motion. They discovered that for attended stimuli, the directional change in a random dot kinematogram produced a larger overshoot in the directional selectivity of the neuronal population, than for the unattended stimulus. This behaviour is consistent with the proposition that attention will lead to a higher precision estimate of new stimuli (Feldman & Friston, 2010) – in this case, the randomly occurring direction change in the kinematograms. However, what Mehrpour et al. (2020) additionally observed was that the directional selectivity of the neurons was increased by more than double when the stimulus was attended than when it was unattended. More specifically, with a directional change of 25°, the directional selectivity of the MT neurons changed to 39° when the change occurred in the attended stimuli, whereas it changed only to 32° when it was unattended. In this way, the authors argue that attention serves the purpose of increasing precision in terms of the importance of a stimulus, favouring a reaction, even if it is an exaggerated one. Perhaps this overshoot facilitates the stimulus to be decoded as substantially different to the preceding stimuli thus initiating a timelier response.

Together these findings emphasise how attention can overturn the saliency of unpredicted stimuli if they were attended and further support the synergic existence of attention and prediction (Feldman & Friston, 2010). Thus, if one wants to be able to distinguish the effects of attention and expectation at the level of both behavioural and neuroscientific results, the two factors need to be explicitly modulated.

1.2.1.3 Critiques of the predictive processing framework

Although the predictive processing framework possesses a lot of explanatory power, it does not come without criticism. In a recently published review by Walsh et al. (2020), the authors argue that the framework itself is very difficult to falsify, as, in essence, it can accommodate almost every outcome. The authors also argue that for a lot of the specific hypotheses such as the existence of distinct neural entities that deal with the predictions and the prediction error, and the hierarchical organisation of the brain there still is not enough supporting research, with studies focusing mainly on the identification of

prediction errors rather than on the formation of the predictions themselves. Despite some of the controversial findings in the literature pointed out by Walsh et al. (2020), their critical review points out that predictive processing provides a framework which could give a unifying architecture for multiple neurocognitive phenomena, including neuropsychological disorders like schizophrenia (Sterzer et al., 2018) and autism (Lawson et al., 2014; Van de Cruys et al., 2014). Moreover, the high scientific interest allows for the development of new testing paradigms which could provide more insight into the predictive processing framework. This refinement of the framework will inevitably lead to a more precise exploration of its assumptions and building blocks, which could lead to further redevelopment (Walsh et al., 2020).

One point of debate around predictive processing is the existence and definition of prediction errors (Walsh et al., 2020). Prediction errors can sometimes be considered as general surprise coming out of neural responses to new stimuli. However, there is a suggestion that repetition suppression (the reduced neural signals and behavioural responses to repeated stimuli) and expectation suppression (the reduced neural activation to expected stimuli as opposed to unexpected) occur on two different time scales, with repetition suppression occurring at the first 40-60ms and expectation suppression later at the 100-200ms (Todorovic & de Lange, 2012). Moreover, underlining the fact that surprising stimuli are not solely responsible for what is termed prediction errors, den Ouden et al. (2009) showed that the brain responds to surprising visual omissions as well. Hence, prediction errors are not simply every new stimulus that occurs – they are not a product of stimulus adaptation - but the prediction error is in relation to prediction. Specifically, den Ouden et al. argue that whereas prediction errors in V1 are in effect the surprise of the presence/absence of a stimulus in the visual field, cognitive prediction errors represent a mismatch at a representational level in higher-order areas.

Another example of controversial findings in the field, directly relevant to this thesis, comes from Garrido et al. (2018). In their task, they instructed participants to attend to one of their ears, while white noise was presented in both ears. The task was to press a button when a gap in the played white noise was present in the attended ear. At the same time, they also incorporated an oddball sound procedure on top of the gap detection task. Garrido et al. (2018)

found that attention increased activation for both the predicted and the unpredicted stimuli arguing against the interaction model between expectation and attention observed by Kok, Rahnev et al. (2012) and theoretically expected under the descriptions of Feldman and Friston (2010). However, in the paradigm by Kok, Rahnev et al. (2012) the predicted and unpredicted stimuli were part of the same paradigm. In this way, they were relevant to the task. On the other hand, what Garrido et al. (2018) modulated in their predicted and unpredicted component of the paradigm was an additional irrelevant distractor stimulus. The attention, which was cued by instructing from which ear the gaps should be reported, did not apply to the sounds, as they were not relevant regardless of whether they were in the attended ear or not. Hence, as the authors themselves comment, a direct comparison cannot be made with the findings from Kok, Rahnev et al. (2012). However, this is an important distinction that needs to be made, as the concept of attention as a way of increasing precision of both predictable and unpredictable signals is that it will aid in the performance of a prediction, or it will minimise entropy. Attention to the sounds would have not aided performance nor minimised entropy, hence the same activation modulation should not be given to stimuli that are not relevant to the task. Moreover, an earlier study by St. John-Saaltink et al. (2015) showed that if the predictable task is irrelevant, expectation suppression is observed for the predicted stimuli, but if the interfering task is taxing, that modulation is not present. Indeed, this study did not modulate attention separately, thus it is not clear how attention would have affected the brain activation in the primary visual cortex. However, what St. Jon-Saaltink et al.'s (2015) findings show is that the relevance of the stimuli to the task, whose predictability is being modulated, matters. These findings are not far from the literature on inattentional blindness in the auditory domain. Inattentional blindness literature argues for this specific event, whereby directing attention to one task would inevitably diminish performance to an unattended task (Jensen et al., 2011). In short, the brain would probably not try to waste energy on stimuli that are irrelevant as it redirects energy sources to the activity that is relevant at present.

1.2.1.4 Summary

In summary, the predictive processing framework provides an appealing explanation of the way people interact with their environment. The framework, although in its development, has seen a large surge of interest, and there is new evidence surfacing constantly that leads to its refinement (Walsh et al., 2020). As evidenced from the reviewed literature above, there has been substantial research that attempts and to some extent succeeds at providing support for some of the major mechanisms on a behavioural and neurocognitive level. There are still avenues that need to be tackled and further support to be provided. However, as it has been suggested even by the critics of the framework (Walsh et al., 2020), predictive processing provides a unifying framework for numerous phenomena which makes it appealing for exploration. One of these phenomena is neuropsychological disorders.

1.2.2 The hypopriors theory of autism or the inability to form predictions about the world

In 2012, Pellicano and Burr published an opinion piece that put forward the Bayesian perspective of autism. Bayesian theories accommodate two components - priors and likelihood. The priors are formed by experience (i.e., the predictions) and allow us to predict the incoming sensory information and to interpret the ambiguity of the environment. The likelihood is the input that is coming into the system from the environment. The combination of the prior with the likelihood then produces a posterior, which attempts to interpret the environment as accurately as possible. Hence, the aim of the prior is to reduce uncertainty. According to Pellicano and Burr (2012), the primary difference in autism is that individuals are not forming appropriate priors about the world. Instead, they form attenuated priors or hypopriors. According to the authors, priors that restrict a person's ability to deal with uncertainty could lead to several problems seen in autism. In fact, problems in dealing with uncertainty have been related to increases of anxiety and have been associated with sensory over- and under-sensitivity (Neil et al., 2016; Wigham et al., 2015).

A broader prior will impede individuals' performance in cases where priors would resolve ambiguity but will improve performance in cases where priors would lead away from the true state of the environment. Therefore, these predictions of

the effect of a hypoprior align with some of the findings in the literature (Pellicano & Burr, 2012). The hypopriors theory provides an account for the reduced susceptibility to illusions, as individuals with ASD will perceive the illusory stimulus as it is since their perceptions will not be affected by prior expectations (Mitchell & Ropar, 2004). The theory also provides a compelling explanation of theories that advocate for cognitive models that directly contrast local and global cognitive styles – i.e., the weak coherence theory (Happé & Frith, 2006) and enhanced perceptual functioning (Mottron et al., 2006). Specifically, hypopriors will reduce the effect of the context and individuals will be able to focus on the individual parts leading to superior performance on tasks such as the Embedded Figures Test, which is commonly used to support these theories (Van der Hallen et al., 2015).

Additionally, the theory attempts to provide an account of the sensory sensitivities observed in autism – both hyper- and hyposensitivity. Due to the uninformative priors that the individuals create, they will not be able to anticipate the incoming sensory information, which would result in hypersensitivities (i.e., extreme aversion to the fire alarm at school) or sensory seeking behaviours as they constantly produce an unexpected experience. When Pellicano and Burr (2012) explain why not all experiences turn into hyper- or hypo-sensitivities, they argue that priors are influenced by the frequency of the occurring events. This premise implies that to create an aversive hypersensitivity, an individual needs to form a specific prior expectation to develop the sensory aversion or sensory seeking behaviours.

Finally, in their opinion paper, they emphasise that Bayesian decision theory provides a tool for explaining non-social features of autism. However, they pose the question of whether social information could also be explained in terms of hypopriors, as social situations are generally complex and ambiguous. They briefly mention the applicability of their account to the social domain by referencing an earlier study where they showed that ASD individuals do not show an adaptation aftereffect to faces (Pellicano et al., 2007). Pellicano and Burr (2012) argue that this effect occurs because autistic individuals are less influenced by priors. An inability to learn from social situations would lead to a lot of uncertainty in social interactions, which could be linked to the high levels of anxiety reported by parents of individuals on the autism spectrum (Neil et al.,

2016). Nevertheless, they do not go into further detail about how their theory would explain the social symptoms.

1.2.3 The theory of inflexible precision given to prediction errors in autism or when every new encounter is important

Pellicano and Burr's (2012) paper sparked conversation around the Bayesian/predictive models of autism. Their opinion piece received numerous commentaries. Brock (2012) argued for too narrow priors instead of too broad ones. Too narrow priors will allow for exact matching on occasions where the prior and environment match, but poor performance in broader contexts. Friston et al. (2013) further elaborated on this suggestion. However, they argued that the predictive processing implementation of Bayesian theory gives a better account of the findings in autism research, where the impact of prior beliefs and sensory information is weighted by the amount of precision (or weight) given to each one. As briefly discussed in section 1.2.1.2, the parameters of these different levels of precision are called hyperparameters and the prior beliefs about them are called hyperpriors (not to be confused with the term hypopriors used by Pellicano and Burr (2012), which they termed less informative priors and is not related to hyperparameters) (Friston et al., 2013). Hence, Friston et al. (2013) argue that in ASD it is the hyperpriors that are improper, rather than the prediction (the prior in Bayesian terms).

Following the conversation around Pellicano and Burr's (2012) paper and the responses around it, in 2014, Van de Cruys et al. proposed the theory of High, Inflexible Precision of Prediction Errors in Autism (HIPPEA). When individuals form a prediction about the environment, their predictions always produce an error signal, because they never exactly match the sensory input (Friston, 2009). Usually, individuals have good hyperparameters and know which prediction errors are worth learning and in what contexts it is more reasonable to rely on prediction, i.e., when the incoming information is too noisy (Friston et al., 2013). According to HIPPEA, ASD individuals are deficient in meta-learning – they do not distribute precision accurately between the prediction and the prediction error. HIPPEA argues that individuals with ASD give higher precision to bottom-up information (the prediction errors) relative to the prediction itself. As prediction errors are the basis for learning, giving too high precision to

prediction errors will lead to learning at every point, forming future predictions that are based on noise and infrequent contingencies. In this sense, the Pellicano and Burr (2012) hypopriors theory and Van de Cruys et al.'s (2014) HIPPEA theory propose almost opposite underlying mechanisms. In a later extension to the original paper, Van de Cruys et al. (2017) focus on the importance of noise and the uncertainty of the environment. They emphasise again that whereas HIPPEA would suggest that inflexibly set precision of prediction errors will inevitably lead to more frequent switching in attempting to learn the new rule, this would not mean that they cannot learn the original rule of association. However, creating a prediction or learning a new rule will take longer to establish. What they argue is that in an unstable environment, when the rule reverses or it changes completely, individuals with ASD should be more willing to switch. They further conceptualise that these individuals cannot distinguish between noise and the volatility of the environment, and this is where the reliance on prediction errors becomes a problem because individuals will not be able to distinguish which new information is relevant and which is just noise.

As mentioned above, accidental variation in the environment will lead to learning that is disconnected from the context. Such learning will lead to predictions that are too rigid and inapplicable in most circumstances. Thus, those predictions will be too specific and will lead to more prediction errors. Similar to Pellicano and Burr's (2012) account, Van de Cruys et al.'s (2014) theory provides an interpretation of findings of exceptional abilities as reported in Mottron et al. (2013). The predictions formed by ASD individuals under the explanation of HIPPEA would allow exact matching, which leads to the superior performance described by Mottron et al. (2013). Exact matching would also explain exceptional and savant abilities commonly reported in ASD (Meilleur et al., 2014). Similarly, research on superior search abilities and the underlying research on enhanced perceptual functioning (Mottron et al., 2006) would be driven by individuals on the autism spectrum allocating more precision to low-level sensory information and the mismatch errors produced from them. Thus, when the task requires an individual to identify a specific figure, prediction errors are highly informative and individuals on the autism spectrum who put higher precision on them will have an advantage when matching the searched

target to the rest of the distractor stimuli. A good example of these is research using local vs global tasks, with the processing style of autistic individuals suggested being driven by local-to-global interference (Van de Cruys et al., 2014; Van der Hallen et al., 2015).

Unlike Pellicano and Burr's (2012) account, HIPPEA provides a better explanation of hyper- and hyposensitivity. Due to the high precision given to prediction errors, accidental events will be learned outside of their 'accidental' context (Van de Cruys et al., 2014, 2019). Chance aversive experiences will build a very inflexible prediction, which will account for future aversion to that stimulus. On the other hand, chance positive experiences will lead to increased sensory seeking behaviour for that positive experience. Both situations would be perceived out of the context of the global situation and thus the positive or aversive experience will be associated with the sensory stimulus. Thus, the requirement for sameness and repetitive behaviours, as well as hyper- and hypo-sensitivity are due to learning, as proposed by Pellicano and Burr (2012), but HIPPEA provides a better theoretical justification. In repetitive environments, where there is little need to estimate the precision of incoming stimuli, the overfitted predictions that individuals have created by interfering in the environments and setting out the routines work well. In this way, individuals on the autism spectrum and their families can minimise the uncertainty about the world by introducing structured environments with low variability.

Finally, unlike Pellicano and Burr's (2012) paper, Van de Cruys et al. (2014) provide a more complete account of social difficulties in ASD. They regard social stimuli as all other types of stimuli encountered in the environment, for which inferences are made continuously. The numerous accidental occurrences in social interactions, where no two situations are the same, along with the rigid predictions and learning of accidental incidents, would lead to the highly pronounced social problems in ASD (Constant et al., 2020). The higher weighting of prediction errors prevents individuals on the autism spectrum to extrapolate more global models about the environment, as uncertainty needs to be additionally modelled for a successful abstraction. However, deeper models about the social world and social interactions require the ability to accurately estimate uncertainty and assign it to different components between the words that someone says, to the context itself, and to the other individual's mental

states. Being unable to estimate where to assign uncertainty - i.e., knowing what needs to be learned, could lead individuals to assign higher level of certainty to sensory information of the situation (Constant et al., 2020; Van de Cruys et al., 2014, 2017). This in turn would lead autistic individuals to develop models about social interactions that are not generalizable. Thus, it would be difficult for autistic individuals to use previous social experiences to guide their present behaviour. Although with sufficient time better and more generalizable models would be formed, it is the case that often these models would have limited applicability (Constant et al., 2020; Van de Cruys et al., 2014, 2017).

1.2.4 The priors versus the prediction errors

Since the conception of the theories described above, substantial research that tests the key assumptions of the hypopriors and HIPPEA theories has been conducted. Firstly, the next section will look into research that investigates autistic individuals' ability to form and rely on priors. In the second section, I will discuss research that investigates whether there is inflexibly high weighting of prediction errors in autism. To do this, I will describe recent research in several fields and will discuss their contribution to the debate.

1.2.4.1 Do autistic individuals form and rely on priors?

One type of paradigm used to investigate the establishment of priors is the adaptation paradigm. It uses continuous presentation of one stimulus, which leads to habituation to the given stimulus. In these types of paradigms, the perception of an ambiguous stimulus is usually further away from the preceding adapting stimulus - i.e., a point-light display (PLD) with a speed profile morphed midway between walking and running, will be most likely perceived as running, if it was preceded by prolonged exposure to a walking PLD. These adaptation effects are described by the tuning of neural populations sensitive to different actions (Webster, 2011). In this example, when presented with the morphed PLD which contains equal amount of information from the walking and running stimulus, the response of the neural population sensitive to the walker will elicit a smaller response than the population sensitive to the running PLD. Thus, the perceptual experience is in favour of the running PLD. In these tasks, the adapting stimulus can be thought of as establishing an expectation and the test

stimulus is then perceived as a prediction error that is perceived further away from the adaptation.

In one such paradigm, Turi et al. (2015) found that children who adapted to a specific number of dots and were then presented with a second display where they had to choose the one with more dots, children on the autism spectrum showed less adaptation and were more accurate at choosing the stimulus that accurately contained more dots. Similarly, van Boxtel et al. (2016) showed that adolescents on the autism spectrum show reduced adaptation to biological motion in a paradigm where they had to categorise an ambiguous PLD into walking or running after adaptation to one or the other. Likewise, autistic children have been observed to not show face identity adaptation aftereffects (Pellicano et al., 2007). Thus, the results in children and adolescents on the autism spectrum seem to consistently show reduced formation of priors, supporting Pellicano and Burr's (2012) account. However, Van de Cruys et al. (2021) showed that children generally form less informative priors irrespective of autism diagnosis status. Moreover, R. Cook et al. (2014) showed that autistic adults do not show a difference from neurotypical adults (NT) in both facial identity and facial expression (angry, happy) adaptation aftereffects. Further, Karaminis et al. (2020) were not able to find differences between autistic and NT adults in speed detection of a PLD after adaptation. However, the differences between Karaminis et al.'s (2020) and van Boxtel et al.'s (2016) findings might come from the type of judgement being made. Whereas in van Boxtel et al.'s (2016) the choice was between a walker or a runner, the identification of whether it is faster or slower might provide a qualitatively different task. Thus, although R. Cook et al. (2014) argue, that the effect might be only present in children and adolescents, this might not be the whole story.

The reliance on priors has also been tested in a variety of other paradigms. For instance, Van de Cruys et al. (2018) aimed to show that individuals with ASD can form and use prior knowledge using Mooney images. To be able to perceive the black and white patches of the images as coherent images, prior information is required – i.e., seeing the original image in colour. In their study, adolescents with ASD did not perform differently from NT individuals. Similarly, in a rubber hand illusion experiment, Palmer et al. (2015) found that on an experiential level, autistic and NT individuals show the same experience of the rubber hand

illusion. However, when the movement of the hand was investigated, NT individuals showed large jerks in their motion, indicating an attempt to correct their gesture between where they expected their hand to be as driven by the illusion and where their somatosensory signals were indicating it was. Autistic individuals, on the other hand, showed no jerk differences with and without the illusion. Thus, individuals on the autism spectrum were more likely to rely on the sensory information that they were receiving, despite the fact that at the same time they were able to create the experience of the illusion as established in the experiment. Incorporating the arguments made by R. Cook et al. (2014) about age, it is possible, that age is not necessarily the main factor, rather the type of task, and the type of priors that are being utilised. In either situation, these findings suggest that autistic individuals can create informative priors, whereas the Pellicano and Burr (2012) account postulates that the priors would be too broad to be informative in these cases.

Other research utilising recently developed priors – i.e., during the experiment - has also shown that individuals on the autism spectrum can learn. A direct comparison between autistic and dyslexic participants in a study by Lieder et al. (2019) showed that ASD participants showed impairment in quick adaptations with a resistance to change but optimal long-term information integration, whereas individuals with dyslexia did not adequately integrate information over time. Further support for being able to form priors during a familiarisation phase is provided by Amoruso et al. (2019). They showed that during the familiarisation phase of a task, where participants observed movement kinematics within specific predictive contexts, children with and without ASD showed no differences in being able to pick the type of the action being observed. Thus, both groups were able to learn the associations between the kinematics, the context, and the outcome. However, when Amorous et al. (2019) changed the length of the videos, classifying this as added noise as it was less informative, they showed that NT children relied more on the established prior than autistic children whose performance remained stable. A similar observation has been made in adults as well by Chambon et al. (2017). In a similar task, they showed that the NT and the autistic individuals all relied on picking the more reliable option rather than the opposing one, and this reliance increased with the increase in noise (reduced length of the video). An additional manipulation that

Chambon et al. (2017) added was whether the action that was being observed was in a social context (two people's hands on the screen) or non-social (only one person's hand). With this manipulation, they observed that the reliance on the prior – i.e., that one action is more likely than another - was larger in the social condition but only for the NT participants. ASD participants, on the other hand, did not show a difference in their reliance on the prior between the two conditions meaning that the social aspect did not differentially affect the performance of the autistic individuals. Van de Cruys et al. (2014) argue that such results need to be interpreted in terms of the social context that they fall in, as the difference that was observed in individuals on the autism spectrum was in the social condition. As mentioned earlier, in social contexts, it is possible that individuals on the autism spectrum are more likely to ascribe uncertainty to the sensory information and thus ignoring the social component (Constant et al., 2020).

From the discussed evidence it appears clear that individuals on the autism spectrum can form predictions/priors about the world and they can learn. However, task variability, the type of priors that the study is testing, and the developmental stage of the autistic individuals are important factors when making generalisable conclusions.

1.2.4.2 Is there inflexible reliance on prediction errors?

From the previous section, it is clear that individuals on the autism spectrum can create and rely on priors. HIPPEA (Van de Cruys et al., 2014, 2017) predicts intact development of priors, albeit slower formation in certain situations and lower-level in comparison to more global and higher-order priors like the ones that would be needed in social situations (e.g. Chambon et al., 2017). HIPPEA's proposition focuses directly on the inflexible precision given to prediction errors (Van de Cruys et al., 2014, 2017). The theory proposes that individuals on the autism spectrum are deficient in meta-learning, meaning that they would not be able to distinguish noisy signals, from those that indicate a change in the environment. Thus, the inflexibly high precision given to prediction errors will lead to learning at every opportunity, leading to faster switching when rules are reversed. Again, the theory emphasises that with longer periods of time and more exposure, development of more generalisable predictions is possible but

these models might still be more low-level than those developed by neurotypical individuals – based on too many details, where matching between predictions and external input is made on lower-level information as opposed to abstract regularities.

Direct results on this topic come from research that focuses specifically on the addition of noise to tasks by making stimuli more ambiguous. One example is a study performed by Pell et al. (2016). In their study individuals with and without autism took part in a task where they had to indicate the eye-gaze direction of a stimulus, with increasing levels of noise. The noise was created by decreasing the contrast of the pupil to the sclera and by adding fractal noise to the eyes. In neurotypicals, the noisier the stimulus, the more likely it is to perceive the eye-gaze to be directed towards the individual, i.e., straight ahead, rather than to one or the other side. It is suggested that when we see others' eyes, they are usually looking at us, therefore with a noisier stimulus, we are more likely to rely on the prior/prediction of direct eye contact. Pell et al. (2016) observed similar magnitude of the effect between autistic and neurotypical individual. Moreover, they also observed that NT individuals with higher autistic traits also did not differ from those with lower autistic traits. Similarly, what is noteworthy in the study by Chambon et al. (2017) discussed in the previous section 1.2.4.1, is the fact that the addition of noise did not seem to show a difference in the performance between individuals. In this sense, individuals on the autism spectrum were able to recognise that the incoming information was noisier, rather than unstable, and relied more on their predictions. This is against the predictions made by Van de Cruys et al. (2017), about the inflexible setting of the precision of prediction errors leading to a reduced ability to disentangle noise from informative signals, as individuals were able to rely more on predictions with the increase in noise. However, this interpretation might be too simplistic. Firstly, the study by Pell et al. (2016) relies on priors that individuals on the autism spectrum could have already developed as HIPPEA argues that with sufficient exposure and time, autistic individuals can develop more generalisable models of the world. Chambon et al.'s (2017) task on the other hand quantified noise as having less information in the used videos. Although that could mean that there is less information present to decide, that does not make the available information noisier as the associations were already

established. Therefore, from these studies, it is difficult to conclude that HIPPEA is not an accurate representation of the results. Thus, to carry on forward, it is necessary for there to be a clear definition of what constitutes as noise in the signal to allow for more specific experimental designs in the future.

Beyond noise, research has investigated prediction errors of autistic compared to NT individuals in situations where they should be modulated based on the consistency of the environment. Using a hierarchical oddball task where participants had to listen to either five identical or four identical and one deviant sound, Goris et al. (2018) observed mismatch negativity (MMN), which is considered to reflect sensory prediction errors. In NT, this is attenuated when the more frequently occurring deviant sound is observed i.e., higher predictability of the oddball sound. Whereas the higher predictability of the oddball produced similar results in autistic and neurotypical individuals, the condition with the lower probability oddball showed that ASD individuals have a smaller MMN signal. This indicates that the autistic individuals are less influenced by the context -i.e., less surprised. This is an important finding, as it indicates two things: Firstly, the absence of a difference in the more common oddball condition between the groups indicates that if the predictability is stable autistic individuals can create predictions of the variability of the environment. Additionally, Goris et al. (2018) did not find a difference in the P3b component in the EEG signal, which is considered to facilitate context maintenance and top-down modulation, although there have been consistent findings of reduced activation in autism (Keehn et al., 2013). Secondly, the smaller effect in the MMN signal in the less common oddball conditions in the ASD group indicates less surprise, which is consistent with findings of the overestimation of environmental volatility in this population (Lawson et al., 2017). If individuals have learned to expect more volatility due to inconsistent nature of the stimulus (Van de Cruys et al., 2017) – 20% of the time as opposed to 80% in the more common oddball condition - then an oddball sound would be less surprising, despite still being surprising. In line with the HIPPEA framework, this result suggests less flexible weighting of the prediction errors in ASD or reduced weighting of prior information. Additionally, a pattern of reduced MMN amplitudes has been observed more commonly in passive tasks such as the task by Goris et al. (2018), thus it might be the task itself that is introducing the

effect. However, reduced MMN amplitudes in passive tasks were most evident in adolescents and not in adults (Y.-T. Chen et al., 2018; Schwartz et al., 2018).

Similar findings of the P3b component have also been observed by Gonzales-Gadea et al. (2015). However, the equivalent activation of P3b was seen to unexpected oddball stimuli – opposite ear, but not for expected ones (same ear). To expected oddball stimuli, individuals on the autism spectrum showed even higher P3b signal than NT, thus indicating heightened top-down modulation. This heightened reaction to expected oddball stimuli could be related to the overestimation of the difference between the expectation and the sensory stimulus (Van de Cruys et al., 2017). On the other hand, expected oddballs showed an equivalent P3a signal, which is related to early deviant detection, but reduced to unexpected oddball stimuli. One important distinction between this task and the task by Goris et al. (2018) is that in the latter the task was passive, whereas the one by Gonzales-Gadea et al. (2015) was an active one. Additionally, the unexpected oddball was unexpected because it was in the opposite ear. This means that the task itself may be at play here, specifically as it has been suggested that active and passive tasks show different results, with active tasks eliminating differences in MMN, and passive ones showing a difference (Dunn et al., 2008; Keehn et al., 2013). What is evident is that there is an imbalance in the way that prediction errors are modulated by top-down information. However, it does not appear to be the case that the precision given to prediction errors is inflexibly high, but high precision is only evident in active tasks.

Finally, as HIPPEA predicts increased switching behaviour due to reliance on prediction errors, a fair amount of literature has attempted to examine the applicability of the predictive coding theories in reward-based paradigms. Recent research by Goris et al. (2019, 2020) showed that reward-based learning provides a different performance profile than simple sensory tasks that do not include a reward component. Specifically, in perceptual tasks – musical preferential tasks with varying predictability of the sounds and a perceptual fluency task in detecting similarity, they found that individuals with higher autistic traits show a preference for the more predictable melody and images, whose contour was present in the priming image (i.e., an expected image). However, within a reward-based gambling task, where one deck was fixed at a

higher winning rate, the participants with higher autistic traits did not show a different preference for the fixed deck compared to the participants with lower autistic traits. Although, individuals with higher autistic traits were faster in choosing the more predictable deck of cards. Thus, whereas there is a higher preference for predictability in individuals with higher autistic traits, it is not directly evident in all contexts – i.e., in reward-based tasks.

The lack of differences in a reward-based task has been found in other studies as well. Manning et al. (2017) set out to test predictions made by the hypopriors and HIPPEA theories directly. In a reward-based experiment, children had to select a pirate chest, which was randomly associated with a reward value of 0:100 points. According to Pellicano and Burr (2012), individuals on the autism spectrum would not benefit from a stable environment where the ratio of the reward between the two chests was kept constant. Whereas according to HIPPEA, in a volatile environment where the ratio of 80:20 switches with 20:80 every 20 trials, autistic individuals should have a higher learning rate. Manning et al. (2017) found no difference between ASD children and neurotypically developing children or adults in the learning rates. Specifically, in the more volatile condition, learning rates were high for both groups. However, in volatile conditions, higher learning rates are more favourable because they would lead to better performance (Crawley et al., 2019). Although the stable conditions are where differences should be observed, the two groups also showed similar performance in the stable condition. These results put both the hypopriors and HIPPEA theories into question. However, as the authors note, the task might have not been challenging enough to make concrete conclusions.

Other researchers have found different results in reward-based tasks. For example, Robic et al. (2015) and Lieder et al. (2019) both found that ASD individuals were more likely to maintain their response for longer, even though a change in behaviour in response to the change in the rule would be more beneficial. Robic et al. (2015) were also not able to find any difference in the stable environment. Similar findings were also observed by Sevgi et al. (2020), where in the stable environment, there was no difference between individuals with high and low autistic traits but when the environment became more volatile, individuals with higher autistic traits did not adjust their learning rate especially when the cue was a social stimulus -i.e., a gaze. In fact, in their

study, individuals with higher autistic traits continued to rely on the gaze even when it was not a good predictor, demonstrating a greater persistence and reliance on the past than on new information. These later findings directly contradict the prediction made by HIPPEA (Van de Cruys et al., 2017) about a higher rate of switching and a greater learning rate due to an inflexible reliance on prediction errors. The results show a greater reliance on the prediction, and a lack of incorporation of the new information particularly in the volatile condition. One possible explanation is that after the reversal, the individuals on the autism spectrum would experience the change in the rules as a completely new situation, which needs to be learned, however, in that case learning should not rely on the prediction created when the previous rule was present.

A later study by Goris et al. (2020) expanded on the reward-based tasks by specifically modulating volatility and noise, by having stable environments with high and low noise and a volatile condition where the 90/10 probability of one set of images providing reward changed every 18 trials. In their paradigm, high noise meant that both images had a 60% reward which was stable for 90 trials, whereas in the low noise, one set of images had a probability of 70% to produce a reward and the other 30%. Goris et al. (2020) found that individuals with more autistic traits perform worse in volatile reward-based environments than neurotypicals but that difference was not evident in stable environments. However, they did not find a correlation with autistic traits in terms of learning rate between the stable environments with low and high noise and the volatile environment. This is consistent with the findings by Manning et al. (2017). These findings also contradict the expectation that individuals on the autism spectrum or high autistic traits would not be able to distinguish between noise and the volatility of the environment (Van de Cruys et al., 2017).

Opposite results have been found by a larger-scale study. In a probabilistic reversal learning task, Crawley et al. (2019) found that individuals on the autism spectrum showed a higher learning rate, which is consistent with the expectations from HIPPEA (Van de Cruys et al., 2014, 2017). In stable environments, children showed a higher learning rate, although a lower learning rate i.e., relying more on past, rather than on recent trial performance would have been optimal. This behaviour was also suggestive of a greater sensitivity to feedback, which is consistent with the higher rates of lose-switch behaviour

across all autism age groups reported in the study. This would be indicative of more reliance on prediction errors in comparison to their prediction as suggested by HIPPEA. Adolescents with ASD, on the other hand, showed a reduced learning rate following reward in comparison to neurotypicals, but no difference in learning rate following punishment, which is in accordance with the additional reduced value sensitivity in this cohort found in the study. This could be suggestive of reduced weighting of positive experiences in adolescents, potentially influenced by the higher rates of lose-switch behaviour found in the study. Finally, in adults Crawley et al. (2019) observed slower updating of behaviour and greater reliance on past experience, but an increased learning rate - a hallmark of the HIPPEA account. Most importantly, behaviours seem to improve with age in all measures, except perseverative errors, which did not show a change in either neurotypical or autistic individuals. Although performance never reaches the performance of the neurotypical individuals, the slower accumulation of evidence in establishing the optimal learning strategy is consistent with HIPPEA.

Overall, it appears that there is no consistent evidence about the higher precision setting of prediction errors in autism. However, as Crawley et al. (2019) suggest, this might be due to differences in the task, as well as differences in the models used to estimate learning rates, along with the smaller samples in general in the autism literature. It is also important to note that HIPPEA suggests that positive and negative experiences can lead to the development of hyper-sensitivities and sensory-seeking behaviours. Thus, it is possible that in reward-based paradigms what we are observing is strong associations between behaviours and positive feedback. In this way, obtaining rewards in reward-based paradigms leads to seeking to stick to the rewarding state and in turn to resistance to change. However, Crawley et al.'s (2019) findings bring that reasoning into question, by showing heightened learning rates after the change, but at the same time also observed sluggish decay. It is then possible that, as Goris et al. (2019) mention, reward-based tasks are simply different than other non-reward based tasks.

Finally, HIPPEA postulates that the effects are caused by high precision setting of the prediction error. As mentioned in sub-section 1.2.1.2, attention is one of the suggested methods through which precision is modulated. However, almost

no research has been done specifically looking at attention in the context of predictive tasks in autism. One recent EEG study has attempted to look at attention and expectation within the same paradigm. Coll et al. (2020) used a hierarchical frequency tagging paradigm, such as the one reported in Section 1.2.1.2. They found that autistic traits modulate the relationship between predictability and intermodulation components in the brain, which are suggested to be indicative of the incorporation of the sensory stimulus and the prediction. The modulation was in such a way that in comparison to participants with higher AQ scores, individuals with lower AQ scores showed a steeper slope for the strength of the signal with increased predictability. In this way, individuals with higher autistic traits needed higher consistency in the predictability of the stimulus to reach a situation where the sensory information is incorporated in the prediction as suggested by HIPPEA. Coll et al. (2020) also found that the integration of low-level and high-level information at higher levels of the hierarchy (the MSPCres slopes) was equivalent for the attended stimuli across the levels of autistic traits. However, individuals with higher autistic traits had a more negative slope in the unattended images than participants with lower autistic traits. Like previous studies suggest, unattended information may be treated differently in the brain for autistic individuals (Orekhova & Stroganova, 2014). It should be noted that the SWIFT result of decrease in the signal-to-noise ratio with increase of predictability was not replicated from the original studies. Thus, although this task and analysis procedure has a high potential to untangle the uncertainties about where in the processing hierarchy differences between autistic and neurotypical individuals occur, the findings by Coll et al. (2020) appear not to support the previously replicated findings and more research is needed surrounding the paradigm. Moreover, it is important to further investigate the role of attention in autism in such predictive paradigms, as active and passive tasks appear to have different effects that can sometimes conflict the findings of the literature.

Based on the research discussed in this section, it is clear that there is evidence for higher weighting of prediction errors in autism. However, that is not the case in all situations. Contradicting findings have been mostly found in reward-based paradigms where higher learning rates and by extension relying more on prediction errors have been observed by some researchers (eg. Crawley et al.,

2019) but not by other (eg. Goris et al., 2020). Additionally, general attention allocation differences in autism (Orekhova & Stroganova, 2014) have the potential to affect the findings and in turn produce conflicting results as seen in passive (Goris et al., 2018) and active tasks (Gonzalez-Gadea et al., 2015). Thus, when interpreting results about inflexibly high precision to prediction errors, the type of tasks needs to be taken into account.

1.2.5 Beyond the hypopriors and inflexible prediction errors

It has been argued that the hypopriors and HIPPEA theories described above might be too simplistic in their explanation of autism (Palmer et al., 2017). In particular, whereas the Pellicano and Burr's (2012) hypopriors theory focuses explicitly on the more simplistic Bayesian model with a focus on the prior, recent studies show that individuals on the autism spectrum can form and rely on priors (R. Cook et al., 2014; Karaminis et al., 2020; Tulver et al., 2019; Turi et al., 2015; Van de Cruys et al., 2018). The HIPPEA (Van de Cruys et al., 2014) theory, on the other hand, takes a more hierarchical approach by emphasising that the difference lies with the precision setting of the prediction error, irrespective of the context but more pronounced in volatile environments. HIPPEA argues that autism is a due to differences in attention allocation and meta-learning (when and where to allocate attentional resources for learning), not a general inability to learn. Thus, in task-switching paradigms where the environment changes from a stable one to a less stable one, or if reversal changes happen in the predictability of a cue, individuals on the autism spectrum will experience the most difficulty.

However, Palmer et al. (2017) and Lawson et al. (2017) both argue that inflexible weighing of the prediction error will lead to learning at every instance – i.e., there will be a higher influence of individual trials and events on the predictive models. Consistent and heightened learning rates, however, do not appear to be an observable characteristic in autism as seen in Section 1.2.4.2. Instead, in the Aberrant precision model of autism Lawson et al. (2014, 2017) and later Palmer et al. (2017) propose that it is an issue with overestimation of the volatility of the environment that causes the high weighting of prediction errors. In this way, if the volatility of the environment is estimated to be high, this will lead to the reduced updating of the predictability of events when

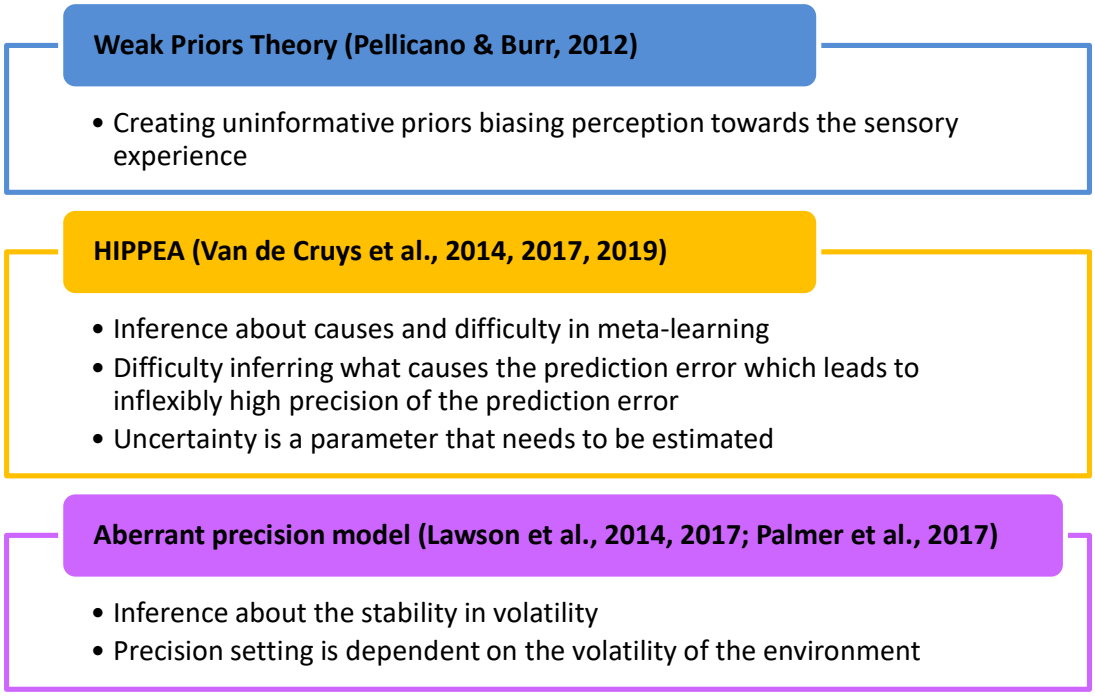
moving to a stable environment. This would make individuals more resistant to changing their estimate about the predictability of an event when the state of the environment changes. This aligns with findings that show that individuals on the autism spectrum show higher resistance to change in the studies by Lieder et al. (2019) and Crawley et al. (2019) discussed in section 1.2.4.2. Palmer et al. (2017) define the difference between Lawson et al.'s (2014) Aberrant precision model of autism and Van de Cruys et al.'s (2014) HIPPEA theory as the difference between inferences about the volatility of the environment and inferences about causes of changes in the environment. Whereas, Lawson et al. (2014, 2017) specifically emphasise the effects of volatility in their research and in defining their Aberrant precision model, Van de Cruys et al. (2014) position their theory with reference to meta-learning - i.e., the inability to learn which prediction errors or infrequent contingencies are to be learned and which ones are to be disregarded. Nevertheless, the main premise of the two theories appears to be the same – the need to know when the volatility of the environment is useful for learning and when it is not. However, since Lawson et al. (2017) found that autistic individuals update their belief about how unstable the environment is more than neurotypicals, it would be reasonable to conclude that the underlying mechanism is not simply about higher precision-weighting of prediction errors as HIPPEA proposes. Moreover, the consistently lower surprise in unpredictable events in individuals on the autism spectrum as compared to neurotypicals is indicative of a more complicated process than an inflexible setting of precision of prediction errors to a higher level (Lawson et al., 2017; Palmer et al., 2017). Whereas these assumptions are reasonable, Lawson et al.'s (2017) findings also suggest that in autism there is a tendency towards overestimation of the stability of cue-outcome associations. This is also purported by Van de Cruys et al. (2017) who argue that not knowing what information about the environment needs to be learned in a frequently changing environment could also lead to learning that the uncertainty is unstable. Specifically, they argue that uncertainty in itself is a learnable state of the world and putting more weight on prediction errors would lead to the same results.

1.2.6 Overall Summary

From what can be seen in Sections 1.2.4 and 1.2.5, prediction errors and the precision associated with them is the more consistent finding in the literature. In

this sense, the argument presented by Pellicano and Burr (2012) about autistic individuals developing uninformative or weak priors does not appear to be supported and the results appear to be more consistent with the HIPPEA theory (Van de Cruys et al., 2014, 2017) and the Aberrant precision model (Lawson et al., 2014, 2017) (See Figure 2 for a summary of the theories).

Figure 2. Summary of discussed Predictive coding/Bayesian theories of autism.



It appears that the premise about meta-learning in HIPPEA can encompass the findings under which the Lawson et al. (2014, 2017) arguments fall. Although Palmer et al. (2017) argue that Van de Cruys et al.'s (2014) theory is about inferring causes and in that way it misrepresents the value of uncertainty, the paper misses out the extent to which the HIPPEA theory emphasises that volatility of the environment will play an important part as a parameter that needs to be estimated. Nevertheless, as mentioned earlier, the mechanism emphasised in both HIPPEA and the aberrant precision model lead to heightened precision of prediction errors in autism. For this reason, the present thesis focuses on investigating whether there is high precision setting of prediction errors.

1.3 Purpose of the present work

As discussed in Section 1.2.1.2, attention is a very important mechanism in the predictive coding framework. It is the proposed mechanism through which precision setting occurs. However, not much research has been done in autism in relation to attention within predictive paradigms, where attention and prediction have been modulated separately. To be able to talk about precision in autism and by extension about the relevance of the HIPPEA theory, attention is a perfect target for research. However, the research discussed so far has focused specifically on learning paradigms and unstable environments.

This thesis aims to contribute research results and paradigm designs that can be used for the investigation of attention in the predictive processing framework of autism. To achieve this aim, this thesis presents three behavioural experiments (lab-based and online), one neuroimaging experiment and one meta-analysis. Each experiment modulates attention and expectation under different experimental paradigms allowing for the investigation of these two factors in more than one context. Additionally, the experiments in this thesis use tasks that utilise moving stimuli. Although, the use of more simplistic stimuli has its advantages, more complex stimuli like motion and biological motion provide dynamics above and beyond simple learning paradigms because they require the integration of low-level and higher-level generalisable concepts. By using motion and biological motion as tools to test the effects of attention in predictive contexts, we add on the effects of pre-existing models about the world.

Firstly, in Chapter 2, I will present an apparent motion paradigm study, which looks at the ability of individuals with and without autism to detect predictable and unpredictable targets that appear on the illusory motion path of two squares that flicker in a frequency that produces apparent motion. This experiment aims to explore the simplistic reality of whether there is higher precision associated with prediction errors in a task that has shown that unpredictable targets are represented in lower visual areas of the brain, but are modulated by higher-order areas that would be responsible for motion perception (Alink et al., 2010; Schwiedrzik et al., 2007). If inflexibly high precision is associated with prediction errors, then individuals on the autism spectrum should be able to detect unpredictable targets better than neurotypical individuals. Moreover, this

experiment controls attention by asking participants to fixate in the centre of the screen, while targets capture attention exogenously.

Next, as social characteristics are some of the most pronounced in autism, I utilise a point-light display of human motion as the next experimental tool to study the HIPPEA theory. Before proceeding with the development of these tasks, however, it was necessary to resolve some of the variability within the autism literature about and whether individuals on the autism spectrum struggle with the perception and interpretation of biological motion stimuli in general. Chapter 3 presents a large-scale meta-analysis of literature investigating the biological motion perception and interpretations in autism in behavioural, eye-tracking, EEG and fMRI studies. At the start of this project, no recent meta-analysis had been conducted on the topic. By the time of publishing, two other groups had started work and published their results at the same time as ours. However, whereas one of these meta-analyses focused on the psychophysics of the tasks that were used (Federici et al., 2020), and the other focused on global motion perception as a whole (Van der Hallen et al., 2019), our meta-analysis had set out to investigate the variability associated with the task at hand and the variability within the participants. We found that there is a general effect of poorer performance in biological motion paradigms for individuals on the autism spectrum, however, this was highly dependent on the age of the participating groups and whether the task was detection, action recognition or emotion recognition (Todorova et al., 2019). Specifically, we found a general decrease in differences between autistic and neurotypical participants with the increase of age, and generally worse performance for tasks that involved emotion recognition, indicating poorer performance with the increase of complexity of information that the participants needed to extract (Todorova et al., 2019). Following these findings from Todorova et al. (2019) presented in Chapter 3, the next two studies focused on using simple biological motion detection in adults, as opposed to emotional or interpersonal action recognition. This allowed us to minimise the complexity of the biological motion stimuli, but at the same time being able to rely on the pre-existing model of human motion.

The second experimental study presented in Chapter 4 orthogonally modulated attention and expectation in two online experiments. This study is a modification of Kok, Rahnev et al.’s (2012) design, transformed to be applied

within a behavioural paradigm, using a biological motion stimulus as an attentional cue. We also developed a second task using coherent dot-motion as a cue to allow the distinction between social and non-social cues. This study aimed to see whether the effects of autistic traits emphasise the effect of attention on unpredictable targets by leading to faster detection of these targets with increase in autistic traits as suggested by HIPPEA.

Next, to better understand the interplay between prediction errors and prediction in autism, Chapter 5 presents an EEG paradigm in which a walking human point-light display was used as a target within a hierarchical frequency tagging paradigm similar to the paradigm presented by Gordon et al. (2017; 2019). Using EEG data, we aimed to see whether there is a difference in the representation of low- and high-level components of the point-light display, along with their integration in the brain of individuals with higher autistic traits as compared to lower autistic traits when varying expectation and attention. The use of moving stimuli in these tagging paradigms has not been done before and thus a proof-of-concept was necessary first. Unfortunately, this study was stopped due to a state of a global pandemic caused by COVID-19 being declared. Thus, only 3 participants' data were recorded. Although limited, the participants' data is provided as a proof-of-concept, and the task should be utilised in the future, as hierarchical frequency tagging presents a promising tool for studying the levels of information integration in the brain.

Finally, Chapter 6 provides a general discussion. The chapter provides a short summary of the findings and how they fit together. It also gives a critical account of the future of the HIPPEA framework and predictive coding frameworks of autism in general and discusses the limitations of the current work giving propositions for further research avenues.

Chapter 2 Special Treatment of Prediction Errors in Autism Spectrum Disorder

2.1 Abstract

For autistic individuals, sensory stimulation can be experienced as overwhelming. Models of predictive coding postulate that cortical mechanisms disamplify predictable information and amplify prediction errors that surpass a defined precision level. In autism, the neuronal processing is putting an inflexibly high precision on prediction errors according to the HIPPEA theory (High, Inflexible Precision of Prediction Errors in Autism). We used an apparent motion paradigm to test this prediction. In apparent motion paradigms, the illusory motion of an object creates a prediction about where and when an internally generated token would be moving along the apparent motion trace. This illusion facilitates the perception of a flashing stimulus (target) appearing in-time with the apparent motion token and is perceived as a predictable event (predictable target). In contrast a flashing stimulus appearing out-of-time with the apparent motion illusion is an unpredictable target that is less often detected even though it produces a prediction error signal. If a prediction error does not surpass a given precision threshold the stimulation event is discounted and therefore less often detected than predictable tokens. In autism, the precision threshold is lower and the same prediction errors (unpredictable target) triggers a detection similar to that of a predictable flash stimulus. To test this hypothesis, we recruited 11 autistic males and 9 neurotypical matched controls. The participants were tasked to detect flashing stimuli placed on an apparent motion trace either in-time or out of time (UT) with the apparent motion illusion. Descriptively, 66% (6/9) neurotypical and (64%) 7/11 autistic participants were better at detecting predictable targets. The prediction established by illusory motion appears to assist autistic and neurotypical individuals equally in the detection of predictable over unpredictable targets. Importantly, 55% (6/11) of autistic participants had faster responses for unpredictable targets, whereas only 22% (2/9) neurotypicals had faster responses to unpredictable compared to predictable targets. Hence, for autistic participants unpredictable targets produce an above threshold prediction error, which leads to faster response. This difference in unpredictable target detection

can be encapsulated under the HIPPEA theory, suggesting that precision setting is aberrant in autistic individuals with respect to prediction errors.

2.2 Introduction

The predictive coding framework argues that the mind is constantly predicting the outcome of the environment, and errors in our predictions are fed back and weighted. Our models about the world are then updated if enough weight is given to the error (Friston, 2010). Based on these models we create future predictions about incoming sensory information, which help us to navigate our environment more efficiently. This is achieved by developing generalisable models of the world allowing for some variability in the input. This is done by appropriately allocating uncertainty to different parts of the environment – i.e., discounting some information as noise, and appropriately emphasizing genuine signals indicating a change (Friston, 2010).

A recent interpretation of autism through the predictive coding framework suggests that autistic individuals develop models that are too narrow, turning small inconsistencies between their prediction and the environment into an error signal (Van de Cruys et al., 2014, 2017). Van de Cruys et al. (2014) argue, that the differences in the neurocognitive processes in autism come from difficulties with meta-learning – knowing when the variability in the environment and its associated uncertainty is a genuine change in the rule and when it should be discounted as noise (Van de Cruys et al., 2014, 2017). Following this line of reasoning, Van de Cruys et a. (2014) propose the theory of High Inflexible Precision of Prediction Errors in Autism (HIPPEA), which postulates that there is higher precision-weighting of prediction errors in individuals on the autism spectrum than in their neurotypical counterparts. Weighing prediction errors consistently high will lead to the development of models that are based on infrequent contingencies due to the noisiness of the environment, creating narrow models of the world.

Individuals on the autism spectrum are consistently reported as having a perceptual style that focuses more on details than the holistic percept (Simmons et al., 2009; Van der Hallen et al., 2015). Having a processing style that focuses on the parts instead of the whole will inevitably facilitate the development of

narrow models through assigning too much weight to sensory information that does not fit the already narrow predictions. Hence, many prediction errors will be registered by individuals on the autism spectrum, which will break down the perception of holistic information (Van de Cruys et al., 2014). Importantly, the HIPPEA theory also argues, that individuals on the autism spectrum can form predictions, but those are often low-level. Unless specifically instructed, autistic individuals will not automatically direct their attention to holistic representations, as low-level features will be easier to predict (Koldewyn et al., 2013; Van de Cruys et al., 2014). One of the proposed mechanisms through which precision-weighting occurs is through attention (Feldman & Friston, 2010) and thus, HIPPEA argues that autism is a disorder of attention allocation.

One way to test whether prediction errors are weighted higher is to utilise an already established paradigm, which rests on a predictive context, where the amount of sensory information - predictable and unpredictable - is varied and attention is controlled. One such paradigm is the apparent motion paradigm as described by Alink et al. (2010). The illusion of motion is created when two identical objects are flickered in rapid succession, thus creating the illusion of a single moving token. This illusory filling-in of the empty frames between the flickering objects impairs the detection of stationary targets shown on the illusory moving token's path (Yantis & Nakama, 1998). This motion masking effect has been shown to vary in strength with the spatial-temporal characteristics of the flashed targets. Targets that appear in-time with the illusory motion token are perceived more readily than those presented out-of-time with it (Alink et al., 2010; Schwiedrzik et al., 2007). The perception of these in-time stimuli invokes smaller activation in the primary visual cortex (V1) which will be expected if the human brain anticipates incoming visual stimuli and thus uses less neural activity to process them (Alink et al., 2010). In contrast, out-of-time targets produce larger V1 activation, which would correspond to stimuli being unexpected, resulting in the brain allocating additional resources to process them (Alink et al., 2010). The behavioural results, whereby predictable (in-time) targets are better detected than unpredictable (out-of-time) ones, indicate that increased activation for the unpredictable targets cannot be attributed to attention and should be viewed as prediction error activation (De-Wit et al., 2010). This is further corroborated by

transcranial magnetic stimulation (TMS) studies. TMS disruption of motion processing brain area V5, before the appearance of the in-time targets, eliminates the advantage in their detection when compared to the out-of-time targets (Vetter et al., 2015). This disruption would be expected under the predictive coding framework as it suggests that higher order areas are responsible for the perception of more holistic stimuli – in this case the perception of motion, and for feeding forward the predictions about where the illusory moving token should be at each point in time.

The apparent motion paradigm has also been tested on schizophrenic individuals (Sanders et al., 2012), which is a condition commonly associated with autism. Sanders et al. (2012) found that schizophrenic patients showed the same advantage in detecting in-time stimuli as neurotypical individuals and greater motion masking than controls – i.e., lower hit rates. These results indicate that schizophrenic individuals were able to form and utilise the prediction created by the apparent motion and show no differences in the processing of prediction errors. Although it has often been suggested that ASD and schizophrenia have a similar underlying mechanism, Van de Cruys et al. (2014) make an important distinction between the two conditions. They argue that the perceptual-cognitive style – local vs global processing, is an underlying reason for their HIPPEA model. This becomes important in light of findings like those from Russel-Smith et al. (2013) where they showed that individuals with high levels of schizotypy have a more global focus, whereas those with high autistic traits have a more local focus. The global focus in neurotypically developing individuals and individuals with high levels of schizotypy would support the similar performance in the two groups in Sanders et al.'s (2012) study. Additionally, a study also looking at autistic and schizotypy traits showed that in a visual statistical learning paradigm, higher autistic traits led to more veridical processing and less influence of expectations, which was due to increased weighting of sensory representations rather than weaker prior formation (Karvelis et al., 2018). This was not true for increase in schizotypy traits. Moreover, Sterzer et al. (2018) put forward the idea that in psychosis/schizophrenia the affected component as explained by the predictive coding framework is the prior. Whereas substantial research is required to tease apart at what levels it is a weak or a stronger prior, the argument that Van de Cruys et al. (2014) are making for autism in HIPPEA is

that in ASD the prediction error is more heavily weighted, rather than having lower/larger weighting on the priors. Thus, we argue that schizophrenic patients should show more 'typical' performance than ASD participants on account of their different processing styles.

It is important to note that it has been argued that individuals on the autism spectrum have difficulty perceiving illusions, although the findings have been contradictory (Simmons et al., 2009). Moreover, it has been shown that the susceptibility to illusions is dependent on the type of illusion used (Ishida et al., 2009). Specifically, David et al. (2010) investigated differences in horizontal and vertical apparent motion perception in ASD. They used the metastable motion quartet, which is a stimulus consisting of two dots alternately presented at four locations – the four corners of a hypothetical square, and thus creating apparent motion illusions in the vertical or horizontal direction. Autistic participants showed reduced horizontal binding in the apparent motion but not reduced vertical binding. Furthermore, individuals on the autism spectrum readily perceive first-order motion - based on luminance (Bertone et al., 2003). Hence, the apparent motion paradigm suggested here should be readily perceived as it is dependent on luminance – i.e., flashing lights with a specific frequency.

Following the discussion above, we propose investigating whether autistic adults show the same advantage in detecting predictable targets over unpredictable ones, as neurotypical participants and whether autistic participants treat prediction errors differently than neurotypicals as proposed by HIPPEA. In this experiment, autistic adults and neurotypical individuals pressed a button every time they detected a target, which was presented either in-time (predictable) or out-of-time (unpredictable) with the illusory motion token's path created by two vertically aligned squares flashing in rapid succession. If prediction errors are more highly weighted in autistic individuals as proposed by HIPPEA, then they will perceive predictable and unpredictable events with the same rate, or unpredictable at a higher rate, whereas neurotypical participants will show an advantage for detecting predictable targets. Thus, as attention appears to be controlled in neurotypical individuals (De-Wit et al., 2010), this task will also allow us to see whether unpredictable targets attract greater levels of covert attention in autistic individuals.

2.3 Materials and Methods

2.3.1 Participants

Twenty participants took part in the present experiment – 9 neurotypical (NT) and 11 autistic (ASD). All participants were biologically male. Participants were group-wise matched on age. To avoid confounds with cross hemisphere communication, we attempted to recruit only right-handed participants, however, two ambidextrous participants on the autism spectrum also took part in the experiment. Participants were also group-wise matched on Full-Scale IQ (FSIQ) as measured by the Wechsler Abbreviated Scale of Intelligence (WASI, Wechsler, 1999). NT participants were screened using the Autism-Spectrum Quotient (AQ) before taking part in the original experiment as it has been shown to provide a good distinction between NT and ASD individuals (Baron-Cohen et al., 2001; Ruzich et al., 2015). A score of 26 was used as a cut off for the NT participants to account for the larger spread of scores in neurotypical populations in Ruzich et al. (2015). Additionally, NT participants needed to have no neurological or clinical/psychiatric conditions/diagnoses. This was not required for ASD individuals as ASD has been found to show high comorbidity with different conditions (Tye et al., 2019). ASD individuals with history of epilepsy were excluded as the stimulation involved flickering stimuli which may risk inducing seizures. From our sample, one autistic participant reported an Anxiety Disorder diagnosis, and another indicated a possible Anxiety Disorder. Additionally, one participant reported cerebellar atrophy of vermis and sulci. They reported that they do not have a problem with motion perception and fixation. They also informed us that they have had surgery to correct for a drifting eye.

All participants' basic vision capabilities were checked for acuity using the Freiburg Vision Test ('FrACT') before any further testing (Bach, 2007). We attempted to confirm the diagnosis of all autistic individuals using the Autism Diagnostic and Observation Schedule (ADOS) by a trained clinical researcher (Lord et al., 2000). Due to the researcher's availability, two individuals did not participate in the interview. The clinical researcher did not provide us with complete scoring from the ADOS and provided us with a binary classification of

the results (would/would not be considered autistic). All individuals reported having received an official autism diagnosis; for nine of those, the diagnosis was confirmed with the ADOS.

2.3.2 Measures

All participants filled in a short demographics and screening questionnaire. It inquired about age, sex, eyesight, and neurological/psychological conditions.

Edinburgh Handedness Inventory (EHI). The EHI (Oldfield, 1971) consist of 10 tasks/activities which involve using one or both hands. Participants are asked to indicate a preference for the right or left hand.

FrACT. The FrACT is an automated, self-paced measurement of visual acuity (Bach, 2007; Brosnan et al., 2012). It consists of a 4-alternate forced choice task, where using a button press participants indicate the orientation of the gap of a Landolt C (contrast 98%), the size of which depends on the correctness of the response. There were 36 trials, where every 6th trial is an 'easy' trial. No auditory feedback was provided. Participants had to have visual acuity of at least 0.5 decimal, which is the legal driving requirement by the Driver & Vehicle Licensing Agency of UK (https://www.gov.uk/driving-eyesight-rules).

AQ. The Autism-Spectrum Quotient was originally developed for investigating autistic traits in individuals. It consists of 50 items to which individuals have to indicate their agreement (Baron-Cohen et al., 2001). NT individuals with scores of or above 26 (Ruzich et al., 2015) were excluded from further testing (N = 3).

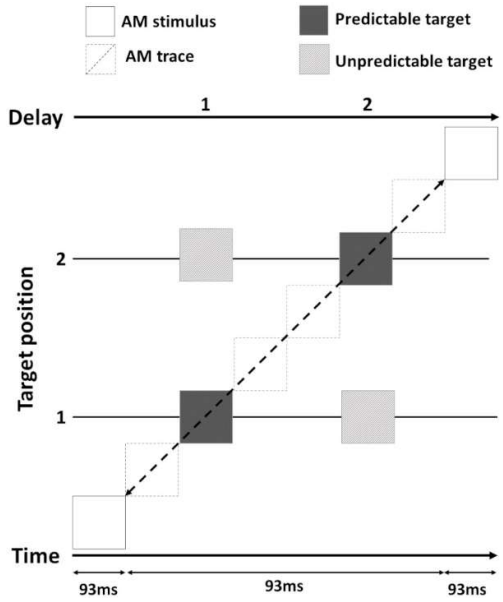
WASI. The WASI (Wechsler, 1999) is an individually administered assessment of intelligence and is applied to individuals aged between 6 and 89 years of age. It provides composite scores of verbal, perceptual and full-scale IQ. It contains four subtests – Vocabulary, Block Design, Similarities, and Matrix reasoning. Both the ASD and NT participants undertook the assessment. As we are looking for individuals of neurotypical IQ, FSIQ scores of below 70 were not included in any further testing. No participants scored below 70. The test was administered either by the ADOS administering researcher or the primary researcher.

ADOS. The ADOS (Lord et al., 2000) is used to assess and diagnose ASD across age and is used for ages between 12 months to adulthood. It takes up to 60 minutes to administer and it consists of semi-/structured tasks, that assess the social and communicative abilities of the individual. A researcher trained to score and administer the ADOS interviewed the participants using Module 4 for adults.

Stimuli. The stimuli replicated the paradigm used by Sanders et al. (2012) with schizophrenic patients. Stimuli were presented using PsychoPy (v1.84) (Peirce et al., 2019) on a CRT monitor (1024x768, 75Hz). The presentation was on a uniform grey background (24.1-29.3cd/m^2)[1] with a white (97.35 – 103.5 cd/m^2) fixation cross (1.2˚x1.2˚) displayed in the centre of the screen. The apparent motion stimuli consisted of two white squares (2.35˚x2.35˚), which flashed to the right of the fixation cross (eccentricity = 7.72˚) and above/below the centre fixation cross (Apparent motion trace = 16.45˚). The stimuli alternated between the two positions, with an inter-stimulus interval equal to the stimuli presentation (7 frames). Presentation frequency was 2.68Hz with one cycle representing one full bi-directional 'motion' - assuming starting from the bottom: $target_{bottom}$(93ms) + ISI(93ms) + $target_{top}$ + ISI. This is more clearly visualised in Figure 3. Targets were white squares (2˚x2˚). Predictable targets appeared in-time with the illusory motion – i.e., a target appearing closer (±4.7˚from centre) to the first flashed stimulus after a short delay will be more predictable than one appearing at a longer delay. As the apparent motion presents a movement of up and then down, this created 4 different target presentations: two in-time (one down and one up) and two out-of-time (one up and one down). Every target was followed by 4-9 apparent motion cycles without a target to maintain the motion illusion. There was a total of 80 predictable and 80 unpredictable targets.

[1] The two values given for the luminance of the background and squares, and the corresponding contrast values (0.559 – 0.603) represent the measurements at lights off and lights on. The two measures were necessary as sometimes the monitor reflected in participants' glasses due to reflective coating and the eye-tracker calibration was not possible in the dark.

Figure 3. Stimuli setup.



***Note*: Predictable targets appeared in time with the motion – i.e., assuming 'motion' starting from the bottom moving up, delay 1 at position 1 and delay 2 at position 2. All stimuli were presented on the right side of the screen. Modified from Sanders et al. (2012)**

Additionally, one control condition was performed by the participants. The control condition used the same stimuli, but it showed simultaneously blinking apparent motion stimuli instead of the temporally displaced ones from the experimental condition. The control condition kept the timing of the targets. In this way, all stimuli appeared at the same locations as before and the targets appeared at the same times, but the illusory motion component was removed by the simultaneously flashing apparent motion stimuli.

2.3.3 Power analysis.

Power calculations were done using the PANGEA (v.02) applet (Westfall, 2016; https://jakewestfall.shinyapps.io/pangea/) as it allows one to take into account the number of replicates i.e., how many trials each participant has in each condition.

For the between-group comparisons, each participant in each motion condition (apparent motion vs no apparent motion) for each predictability level (predictable vs unpredictable) will see 80 targets i.e., 80 trials. The sample size calculated for the interaction of interest (Group x Condition x Predictability) was performed with an estimated effect size of d = 0.45 as recommended by Westfall (http://jakewestfall.org/publications/pangea.pdf, 2016) as no other estimate is

available. The recommendation is based on the meta-analysis of meta-analyses in social psychology by Richard, Bond Jr., and Stokes-Zoota (2003). We were not able to find another study that compared an autistic and a neurotypically developing population on a similar task. Additionally, although Sanders et al. (2012) have used the same paradigm to study schizophrenia, the effect size needed was not reported, because the group interaction was not significant i.e., Schizophrenic and neurotypical volunteers showed the same predictability effect. Moreover, Sterzer et al.(2018) and Van de Cruys et al. (2014) call for more research to distinguish between the two conditions, hence, it is not appropriate to generalise between these two conditions. Thus, using a moderate effect size (d = 0.45) as a guideline, we calculated that with 20 participants in each group we will reach power above 0.95.

Nevertheless, using a smaller effect size of 0.33 found in a meta-analysis of overall coherent motion perception difference between ASD and neurotypically developing individuals (Van der Hallen et al., 2019), we found that with the same number of participants (20 per group) we will still reach a power of ~0.93.

In terms of replicating previous findings in neurotypical participants, the effect of condition consistently shows strong effects - i.e., η^2= 0.853/d = 4.8 [N = 8] (Schwiedrzik et al., 2007); and η^2 = 0.565/d = 2.27 [N = 31 per group] (Sanders et al., 2012). For the interaction condition*predictability, effect sizes have also been consistently large – i.e., η^2 = 0.952/d = 8.9069 [N = 8, condition*delay*position interaction or position*delay at the apparent motion condition only η^2 = 0.816/d = 4.2118 (Schwiedrzik et al., 2007); η^2= 0.067/d = 0.536 for condition*predictability*position [N=31 or η^2 = 0.109/d = 0.6995 for condition*predictability interaction] (Sanders et al., 2012). When decomposing into simple effects, the predictability effect in the apparent motion condition has been found to be d = 1.086 [N = 8, only apparent motion condition, one target position (bottom)] in Edwards et al. (2017) and d = 0.494 [N = 31, only apparent motion condition for neurotypical participants, two target positions (top & bottom)] (Sanders et al., 2012). The smallest effect size of 0.494 suggests that our initial sample size of 20 participants per group would only reach power of 0.719. To reach power of at least .80, a minimum of 25 participants will be required.

Thus, the sample size for both groups was chosen to be 25 to ensure that we could detect the effect in the neurotypically developing group. However, due to the COVID-19 pandemic and the national lockdown in Scotland associated with it, the study was stopped before the full sample could be reached, and we were only able to recruit 11 autistic and 9 non-autistic individuals. Additionally, due to the use of the eye-tracker, the ADOS, the WASI and the visual acuity test, the experiment could not be transferred online. Therefore, we are publishing this work along with all analysis and experiment set up scripts to allow other groups to build up on the data we have collected.

2.3.4 Procedure

Participants performed the FrACT, EHI, AQ, WASI before the beginning of the experiment. Participants took part in the ADOS before or after the experimental task depending on participant/clinical researcher availability. If participants did not meet inclusion criteria on the AQ, WASI or FrACT they did not continue with further testing (n=3).

Afterwards, we seated the participants at a viewing distance of 70 cm and asked them to fixate on the fixation cross at the centre of the screen throughout the experiment, while allowing for brakes in fixation between trials. The use of the eye-tracker allowed us to monitor participants' fixations and they were told if their eyes started to drift to perform a corrective eye-movement. Additionally, the eye-tracker allowed us to monitor for blinks, thus removing trials where the target appeared at the same time as a blink occurred. We monitored the fixations by tracking eye-movements of the right eye of the participant throughout the experiment using a remote eye-tracker EyeLink 1000 v4.51 at a 250Hz sample rate. We presented one block per condition (experimental/control), and we counterbalanced the order across participants. We calibrated the eye-tracker with a standard 9-point calibration with each subject. The participants maintained their head-position throughout the experiment with the help of a chin rest. Between the blocks, we encouraged the participants to take a break and rest their eyes. We instructed participants that they will see a square appearing at the top and at the bottom of the right side of the screen, and that they needed to press a button every time they saw a square

appear in between those two locations. Participants performed a short practice of 12 target presentations at the start of each block. We asked the participants to respond as quickly as possible to the perception of the target by pressing the space bar on a standard keyboard. A total of 160 targets were presented across the two blocks with a 30s break after every 15th target where participants free-viewed nature scenes presented on the screen, without moving from the chin rest. Every response between 150 and ~1500ms (4 apparent motion cycles) after target presentation was considered as a hit and everything else as an erroneous response. Each block lasted approximately 15 min.

2.3.5 Analysis

The analysis was carried out in R(v4.0.4) (R Core Team, 2020) using RStudio(v1.3.1093) (RStudio Team, 2016) using tidyverse (v1.3.0) and attached packages (Wickham et al., 2019).

For a target to be coded as detected the keypress should have occurred between 150ms and 1500ms after the target onset. This response window was selected based on previous research (Sanders et al., 2012; Schwiedrzik et al., 2007).

Trials, where the participant performed a blink when the target was presented were excluded (N=43 across participants (apparent motion=24, control=19)). Additionally, 90 more trials were removed (all from the ASD group) where participants either misunderstood the task or there were problems with the eye-tracking equipment. Despite reporting that they have had corrective surgery for their eye, the participant that reported cerebellar atrophy had large shifts in their eye movements, for which they performed consistent corrective eye-movements. For two additional participants, eye-tracker calibration was not possible (1- both condition, 1- apparent motion condition only) and a verbal reminder was given about maintaining fixation at the beginning of each set of 15 trials (after the break). For that reason, their eye-movement data is not included in the descriptive analysis below in Table 1.

Due to the small sample size, the data are presented and interpreted descriptively. For completeness, binomial regression (for performance data) and gamma regression (for reaction time (RT) data) were also used for fitting the data, without the inclusion of random intercepts and slopes. The initial pre-registered analysis resulted in many convergence failures resulting from too many parameters and not enough data points. The analysis here deviates from the pre-registered one, to provide complete description of the data. The original pre-registered analysis can be found at the project folder at OSF (https://osf.io/avsqh/) and analysis scripts and data will be uploaded on ReShare. Significance of all effects was done through model comparison of the full model, with a reduced model that excluded the effect of interest.

Initially, to check if we could replicate previous findings in neurotypically developing individuals and to confirm the paradigm performed as expected, we ran a logistic regression with condition (apparent motion vs no apparent motion) and predictability (predictable vs unpredictable) for only the NT group.

Next, to directly test the hypotheses, logistic regression was used to check for the 3-way interaction between condition (apparent motion vs no apparent motion), predictability (predictable vs unpredictable) and group (ASD vs NT). Finally, the ASD group's performance only was analysed to investigate any potential differences in the control/apparent motion condition.

As an exploratory measure, reaction time was also explored to tap into any processing delays within the ASD population. An analysis was run with reaction times (for the detected targets) as the outcome variable rather than the responses.

Finally, we investigated whether Age and FSIQ were significant covariates to the model.

2.3.6 Ethics

Ethical approval was obtained from the Ethics Committee at the College of Science and Engineering, University of Glasgow. All participants provided informed consent for taking part in all parts of the study and for their

anonymised data to be shared. All participants were allowed to take a copy of their scores.

2.4 Results

2.4.1 Descriptive Statistics

As seen in Table 1, the two groups were not significantly different on age or FSIQ but were significantly different in AQ scores. Percentage of fixations within the 2° window away from the centre of the screen were compared between the groups for each condition. Comparisons were performed using Welch's Two Sample t-tests.

Table 1. Descriptive Sample Statistics

Group	ASD	NT	*t*	*df*	*p*
Mean Age (SD)	33.73 (13.84)	29.78 (12.74)	0.66	17.71	0.52
Mean AQ (SD)	35.27 (7.34)	12.00 (5.50)	8.10	17.9	<0.001
Mean FSIQ (SD)	115 (14.90)	118.44 (12.61)	-0.56	17.96	0.58
Mean % fixations out of centre, apparent motion (SD)	15.50 (14.82)	4.96 (3.21)	1.97	7.59	0.09
Mean % fixations out of centre, control (SD)	6.83 (7.16)	4.7 (2.75)	0.83	10.31	0.42

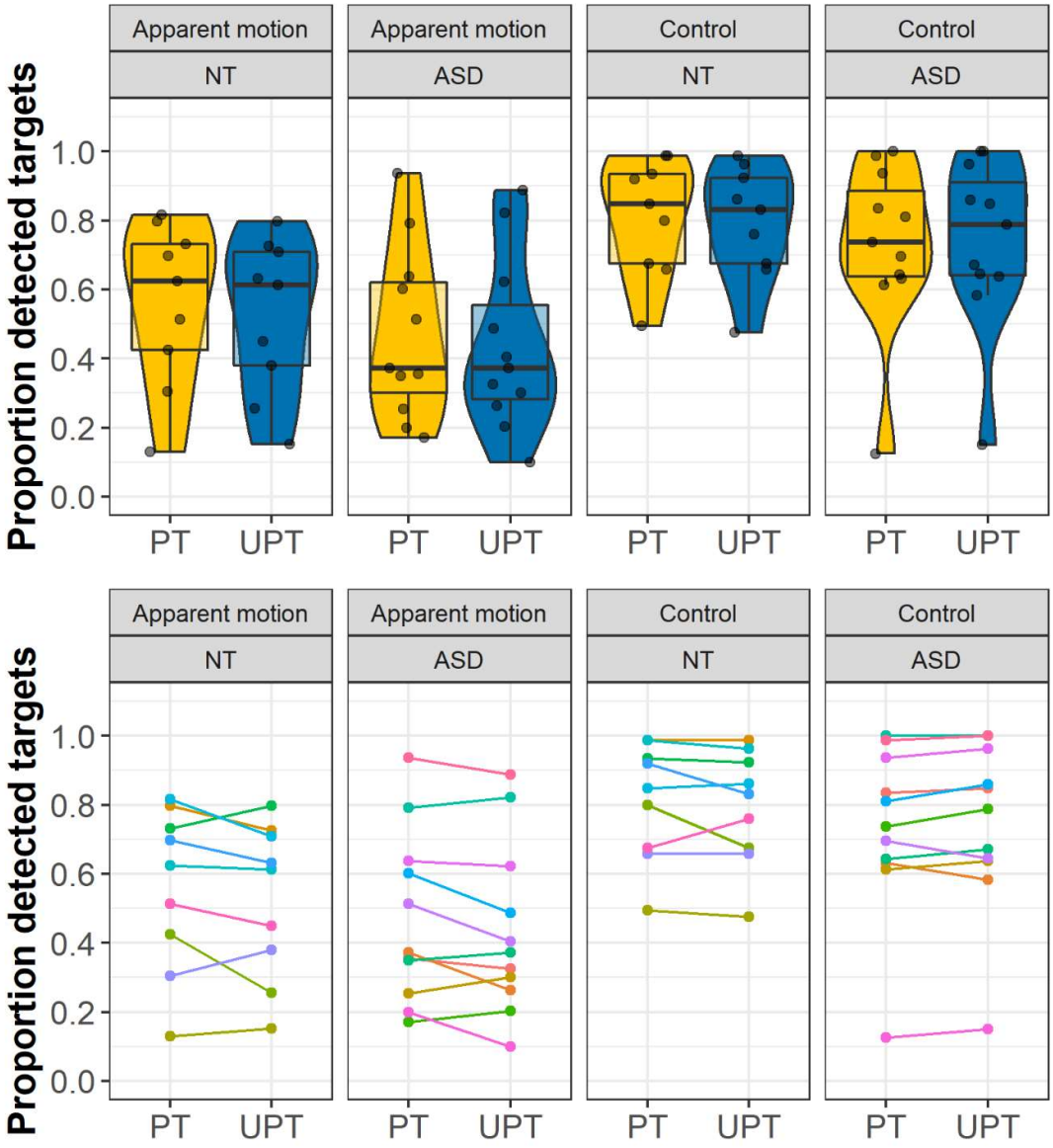
Table 2 describes the participants' performance based on condition and predictability. Overall, it appears that the control condition was easier than the apparent motion condition, which indicates that the apparent motion introduced motion masking (see Figure 4). Additionally, the NT participants performed better in both conditions. From Table 2 it appears that performance was better for the predictable than the unpredictable targets in the apparent motion condition for both groups, with 66% (6/9) of NT participants and 64% (7/11) of ASD participants detecting more predictable than unpredictable targets. This can more clearly be seen in the lower panel of Figure 4.

Table 2. Participants' performance by condition and target type for each group

Condition	Target	Group	Mean detection rate (SD)	Median reaction time (ms) (MAD*)
Apparent Motion	Predictable	NT	0.56 (0.5)	478 (152)
		ASD	0.47 (0.5)	510 (200)
	Unpredictable	NT	0.52 (0.5)	519 (157)
		ASD	0.44 (0.5)	501 (180)
Control	Predictable	NT	0.81 (0.39)	418 (83)
		ASD	0.73 (0.44)	416 (95)
	Unpredictable	NT	0.79 (0.41)	420 (91)
		ASD	0.74 (0.44)	418 (101)

***Note*: *MAD – Median Absolute Deviation**

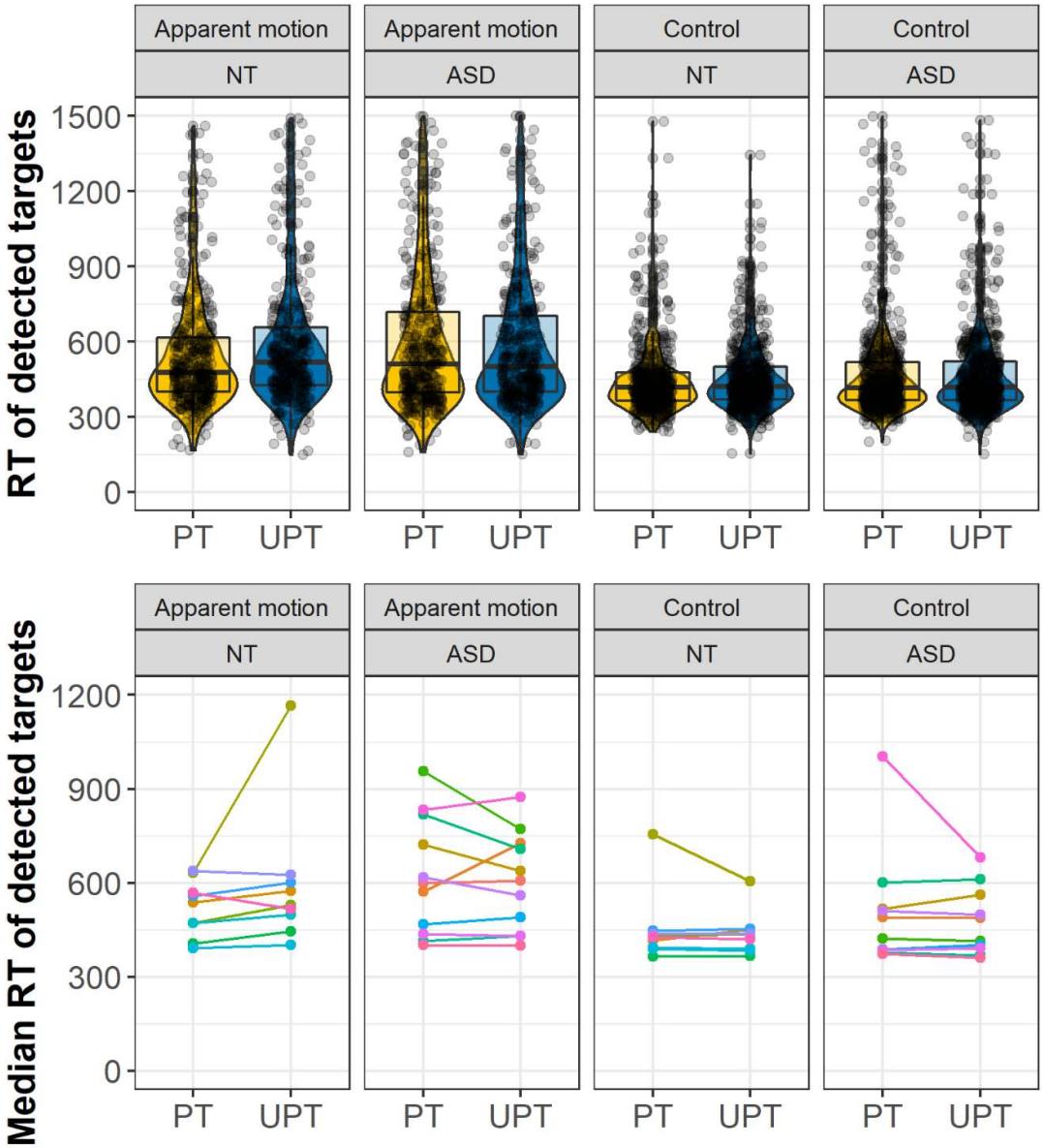
Figure 4. Average performance per person for each group per condition.



***Note*: PT - Predictable Target; UPT - Unpredictable Target. Top panel present the overall performance per group with individual data points for each participant. Bottom panel shows the data as pairs of observations for each participant between target conditions**

Looking at the RTs, from Figure 5 and Table 2 we can see that NT individuals were slower at detecting unpredictable targets, whereas ASD participants were faster. The lower panel of Figure 5 more specifically highlights that while 22% (2/9) of NT participants had shorter median reaction times for the unpredictable targets, more than half of ASD participants (55% - 6/11) had shorter median reaction times, indicating faster detection of unpredictable targets in the ASD group. It is noteworthy that the variance is rather large for both the performance and RTs, thus these observations should be taken with caution.

Figure 5. Reaction time (ms) for each group per condition



***Note*: PT - Predictable Target; UPT - Unpredictable Target. Top panel presents violin and boxplots of reaction time per target and condition with reaction time point superimposed for each participant and trial for each condition. Bottom panel shows the paired median RT for each individual between target conditions.**

2.4.2 Performance analysis

2.4.2.1 Replicating previous findings in neurotypicals

In the first instance, the effects only in the neurotypical population were analysed. A comparison between the full model and model without the main effect of condition (apparent motion vs control) showed that there was a significant effect of condition – $X^2(1) = 220.34$, $p < 0.001$. NT participants were able to detect a larger proportion of targets in the control condition ($M_{Apparent\ motion}$ =0.540 [SD:0.499], $M_{Control}$ =0.801 [0.399]) where there was no apparent motion illusion present. Unfortunately, the interaction effect between condition and predictability was not replicated – $X^2(1) = 0.02$, $p = 0.88$. As the simple effect of predictability would not provide us with any information because it will be across conditions, the effect was not explored.

2.4.2.2 Overall analysis

There was an overall effect of group, indicating that participants on the autism spectrum were less likely to detect the targets overall (M_{ASD} = 0.596 [0.491], M_{NT} = 0.672 [0.470]) - $X^2(1) = 29.26$, $p < 0.001$. There was no effect of the group and condition interaction - $X^2(1) = 0.03$, $p = 0.86$. This is also evident from Table 2. Additionally, there was no three-way interaction between group, condition, and predictability - $X^2(1) = 0.65$, $p = 0.42$. No further effects were looked at.

2.4.2.3 Performance in the ASD group only

Finally, as there was not a significant effect in the neurotypically developing group for predictability, but most participants detected a higher proportion of predictable than unpredictable targets, to maximise power, and to establish any effects of predictability in the apparent motion condition for ASD participants only, we ran a model that tested the simple effects of predictability for ASD participants in the apparent motion condition. However, the simple effect of predictability was not significant when comparing the modified model, and the same model without the predictability variable - $X^2(2) = 2.27$, $p = 0.32$.

2.4.2.4 Covariates

There was no significant effect of age in the omnibus model - $X^2(1) = 0.06$, $p = 0.81$. However, FSIQ was a significant covariate - $X^2(1) = 72.21$, $p < 0.001$. This indicates that higher IQ was to an extent responsible for higher detection rates.

2.4.3 Reaction Times

2.4.3.1 Overall analysis

There was an overall effect of group, which was driven by NT having faster reaction times on average (M_{NT} = 498 [198] (median:442 [114]), M_{ASD} = 517 [228] (442 [126])) - $X^2(1) = 1.91$, $p = 0.01$. There was a significant effect of condition - $X^2(1) = 45.58$, $p < 0.001$, indicating faster detection in the control condition ($M_{Apparent\ motion}$ = 579 [257] (502 [174]), $M_{Control}$ = 464 [169] (418 [92])). This is also evident from Table 2. The group and condition interaction was not significant - $X^2(1) = 0.05$, $p = 0.56$. Additionally, there was no three-way interaction between group, condition, and predictability - $X^2(1) = 0.31$, $p = 0.15$. No further effects were explored.

2.4.3.2 Covariates

There was not a significant effect of age at the omnibus model - $X^2(1) = 0.50$, $p = 0.07$. However, FSIQ was a significant covariate - $X^2(1) = 9.36$, $p < 0.001$. This indicates that higher IQ of individuals was to an extent responsible for faster reaction times.

2.5 Discussion

This experiment aimed to investigate whether predictable and unpredictable targets are detected differently in an apparent motion paradigm between individuals with and without autism. According to the HIPPEA theory proposed by Van de Cruys et al. (2014), there is high inflexible precision of prediction errors higher in autism. Therefore, unpredictable targets should be more easily detected by autistic individuals. The results observed in the present experiment show that the illusory motion made the task more difficult for both NT and ASD groups. This indicates that both groups were able to experience motion masking. However, there were no effects of predictability in either group and there were

no interactions with group on performance. The only effects of group showed that autistic individuals detected fewer targets and were slower in their responses. However, a descriptive examination of the data shows that precision for prediction errors as represented by unpredictable targets is set differently in autistic individuals than in neurotypicals.

The descriptive results, although incomplete, highlight some important trends in the data. By count, it is evident that most of the participants in both groups detect more predictable than unpredictable targets. This supports HIPPEA's argument, that individuals on the autism spectrum can form predictions. In the present case, the formation of a prediction about the 'movement' of the token facilitated performance in both groups for the predictable targets and led to a decreased ability in detecting unpredictable targets. These results mirror the findings from Alink et al. (2010) and Schwiedrzik et al. (2007).

So far, the results are similar to the finding by Sanders et al. (2012) with schizophrenic patients. However, in the present study autistic participants additionally showed a descriptively faster detection for the unpredictable targets, which could be indicative of the higher precision associated with prediction errors as suggested by HIPPEA (Van de Cruys et al., 2014). As Feldman and Friston (2010) emphasise, attention will have a spotlight effect on prediction errors, assisting in their propagation up the processing hierarchy. Thus, although participants on the autism spectrum still experience the motion masking effect created by the apparent motion paradigm, and the prediction appears to assist them with the detection of predictable over unpredictable targets, unpredictable targets appear to be given special treatment. Moreover, the fact that overt attention is not modulated in this paradigm (De-Wit et al., 2010) suggests that attention might be disproportionately affecting prediction errors, capturing the covert attention of autistic individuals to a higher degree, leading to faster reaction times.

There are some limitations to this study that need to be considered when moving forward. The sample size used here is not appropriate as shown by the power calculation. Thus, an appropriate sample size should be recruited and the analysis should use design-appropriate models as proposed in the pre-registration. Above and beyond the insufficient sample size, we only recruited

male participants to avoid any interaction effects with sex. It has been shown that brains of autistic individuals have differing connectivity between the sexes and in comparison to their neurotypical counterparts (Alaerts et al., 2016; Lawrence et al., 2020). However, to be able to characterise the complete ASD profile, we need to know whether HIPPEA can explain differences in both sexes. Samples with only male, only female and comparison between the two are necessary to achieve this. Thus, we are making available the analysis and experiment scripts along with the pre-registration of the analysis for future researchers to add to this dataset. Datasets from multiple sites with diverse samples will allow for the establishment of more robust findings. This in turn will allow for future research to have a clearer path moving forward when testing HIPPEA's predictions.

Further, it is noteworthy that there is research suggesting that autistic individuals tend to use compensatory brain networks to show similar behavioural responses (McKay et al., 2012; Philip et al., 2012). Therefore, our sample of autistic males might show similar behavioural performance to the neurotypical population because they are using compensatory brain networks. Thus, this task may be too simplistic to show differences in the ability of individuals with autism to form and utilise predictions to guide their behaviour. Thus, apart from adding more data points to this study, it will be important to consider more complex stimuli which also control for attention and expectation at the same time.

The present results weakly suggest that in autism there is special treatment of prediction errors as expected under HIPPEA. However, it is important to point out that the formation of predictions as seen by the descriptively better detection of predictable targets along with the motion masking effects suggest that our autistic participants were able to form predictions, which is one of the pillars of HIPPEA. These results are promising and the recruitment of a sufficient number of participants, as suggested by the power analysis, will be necessary to establish how significantly reliable these results are. Nevertheless, the findings take us one step closer to finding out whether HIPPEA is a good candidate for explaining autism.

Chapter 3 Biological Motion Perception in Autism Spectrum Disorder: A Meta-analysis[2]

3.1 Abstract

Background: Biological motion, namely the movement of others, conveys information that allows the identification of affective states and intentions. This makes it an important avenue of research in autism spectrum disorder where social functioning is one of the main areas of difficulty. We aimed to create a quantitative summary of previous findings and investigate potential factors, which could explain the variable results found in the literature investigating biological motion perception in autism.

Methods: A search from five electronic databases yielded 52 papers eligible for a quantitative summarisation, including behavioural, eye-tracking, electroencephalography and functional magnetic resonance imaging studies.

Results: Using a three-level random effects meta-analytic approach, we found that individuals with autism generally showed decreased performance in perception and interpretation of biological motion. Results additionally suggest decreased performance when higher order information, such as emotion, is required. Moreover, with the increase of age, the difference between autistic and neurotypical individuals decreases, with children showing the largest effect size overall.

Conclusion: We highlight the need for methodological standards and clear distinctions between the age groups and paradigms utilised when trying to interpret differences between the two populations.

Keywords: autism spectrum disorders, biological motion, meta-analysis, age, emotion recognition

[2] This chapter was published in Molecular Autism: https://doi.org/10.1186/s13229-019-0299-8

3.2 Introduction

Biological motion (BM), namely the movement of other humans, conveys information that allows the identification of affective states and intentions (Blake & Shiffrar, 2007; Pavlova, 2012; Pollick et al., 2001). BM processing specifically, is the ability of individuals to detect, label and interpret human movement and to allocate certain emotional states to it. Thus, BM is an important component of social perception. Moreover, neurotypically developing (NT) individuals have been shown to be able to readily extract socially relevant information from sparse visual displays (Blake & Shiffrar, 2007; Pavlova, 2012). Specifically, point-light displays (PLDs), which portray BM with points located only on the major joints, are readily recognised as depicting differing actions by NT (Johansson, 1973).

Pavlova (2012) argues that an inability to extract socially relevant information from BM could have damaging effects on social functioning. In fact, individuals with an intellectual disability have been shown to have no problem in identifying different types of motion (Klin et al., 2009; Sparrow et al., 1999), whereas individuals with social functioning difficulties such as Autism Spectrum Disorder (ASD), have shown reduced ability in extracting social information from BM (Simmons et al., 2009). Indeed, ASD's main diagnostic characteristics include problems with social interaction and communication as well as repetitive and/or restrictive behaviours (American Psychiatric Association, 2013). Thus, the social impairment in ASD can, to some extent, be readily related to a reduced ability to extract information from BM.

However, findings on BM in ASD tend to be mixed (Simmons et al., 2009). For example, some studies, which investigated the identification or recognition of actions from BM (Hubert et al., 2007; P. Murphy et al., 2009; Parron et al., 2008; Saygin et al., 2010) did not find significant differences between NT and ASD individuals, whereas others have found differences between the two groups (Annaz et al., 2010; McKay et al., 2012; Nackaerts et al., 2012). Simmons et al. (2009) and McKay et al. (2012) argue that this is because there is variability between ASD individuals. Several factors have been suggested to introduce this variability.

One of these potential factors is age. Specifically, on one hand it appears that research in children tends to consistently show an impairment in BM interpretation (Annaz et al., 2010; Blake et al., 2003; Klin et al., 2009). Whilst, on the other hand, research in adults does not find differences in performance in action perception and BM recognition (Hubert et al., 2007; P. Murphy et al., 2009; Saygin et al., 2010).

Person characteristics such as sex and IQ have also been suggested to contribute to the variability of results. Specifically, IQ has been identified as a predictor of performance in some studies (Jones, Swettenham, et al., 2011; Rutherford & Troje, 2012) but not in others (A. P. Atkinson, 2009; Hubert et al., 2007; Van Boxtel et al., 2016). Furthermore, a recent meta-analysis by Van der Hallen et al. (2015) looked at local vs global paradigms, where individuals have to ignore the global context to be able to focus and perform a task on the specific parts or vice-a-versa. They observed greater differences when the proportion of females was higher. Hence, these demographic characteristics of the samples should be investigated as potential contributors to the variability in the findings.

The task at hand has also been considered as a contributing factor. Koldewyn et al. (2010) argue that individuals with ASD are able to identify BM presented through simple PLDs from noise and classify them, however, it is the extraction of higher order information, such as emotional content, that shows the largest performance difference. In fact, although Hubert et al. (2007) and Parron et al. (2008) did not find differences between NT and ASD in action recognition, they found differences in emotion recognition from biological motion for adults and children. Additionally, Fridenson-Hayo et al. (2016) found that in children this difference in emotion recognition from BM is evident for both basic (e.g. happy, sad) and complex emotions (e.g. disappointed, proud) as well as being evident cross culturally (Britain, Sweden, Israel). Thus, both children and adults with ASD tend to be less sensitive to emotional content.

It has been suggested that eye-tracking research can inform our understanding of the social difficulties in ASD. A review and meta-analysis of eye-tracking studies showed that in ASD attention to social versus non-social stimuli may be reduced (Chita-Tegmark, 2016). The analysis also found that decreased attention

might be given to the eyes and increased attention to the mouth and body compared to NT individuals. However, Chita-Tegmark (2016) noted that the results were very mixed. This may have been because the authors tried to include a large number of studies and thus inevitably included a mixture of more than one type of stimuli, including faces, eyes and bodies. Specifically, bodies contain vital social information and are perceptually different from faces (de Gelder, 2009). Thus, different processes may be involved when looking at these different stimuli. Nevertheless, even when looking at eye-tracking studies focusing only on biological motion, the same variability is observed. Namely, in preferential looking paradigms, children have shown reduced visual orientation to biological motion (Annaz et al., 2012; Falck-Ytter et al., 2013; Klin et al., 2009). This difference between NT and ASD has not been found in adults (Fujioka et al., 2016). In contrast, Fujisawa et al. (2014) show that pre-school children tend to have a greater preference for upright than inverted BM, which was additionally greater than that in NT children. Hence, it is apparent that inconsistencies in eye-tracking studies also exist but cannot be simply explained by age as a driving factor.

One study argued that the mixed findings in the BM literature within ASD are due to ASD utilising different brain networks which develop later in life. Hence, McKay et al. (2012) investigated BM perception between ASD and NT and found that the brain areas that communicate with each other in ASD are not the same as the ones found in NT. Specifically, functional magnetic resonance imaging (fMRI) studies tend to find reduced activation in ASD for areas such as the superior temporal sulcus, middle temporal gyrus, inferior parietal lobule. These are all areas that have been found to be related to the perception and interpretation of human motion and actions (Alaerts et al., 2017; Freitag et al., 2008; Grèzes et al., 2009). NT individuals, however, show connectivity within areas involved with action and human motion observation – such as the inferior and superior parietal lobules. On the other hand, individuals with autism have been found to have brain networks that involve connectivity with the fusiform, middle temporal and occipital gyri, which are all areas considered to be involved in more basic level motion perception rather than action recognition (Freitag et al., 2008; McKay et al., 2012).

Similarly, the mirror neuron network (MNN) has been implied to be related to social functioning as it is associated with observing and understanding the actions of others. Thus, Kaiser and Shiffrar (2009) argue that the MNN could contribute to the impairments seen in ASD. Moreover, Villalobos et al. (2005) have shown reduced functional connectivity in the prefrontal mirror neuron area in individuals with ASD. The MNN has mainly been investigated in imitation paradigms (Oberman et al., 2005; Williams et al., 2006) and indeed, dysfunctional activation has been identified in individuals with ASD. However, since the MNN is also involved in understanding others' actions, its activation during simple action observation has also been investigated in ASD because understanding others' actions is an integral part of social functioning. Most commonly, mu-suppression has been used to assess human mirror activity (Fox et al., 2016) and reduced mu-suppression has been found in ASD participants in comparison to NT individuals both when performing and observing BM (Oberman et al., 2005; Raymaekers et al., 2009). Thus, it appears that the impairment in the MNN could be another contributing factor to the social difficulty present in BM perception in ASD.

In order to help bring clarity to the field, there is a need for a quantitative review of the research done on BM perception in ASD. Previous literature reviews have already argued for reduced ability in interpreting social information from BM and about the diagnostic utility of biological motion in ASD (Kaiser & Pelphrey, 2012; Kaiser & Shiffrar, 2009). In one such attempt, Van der Hallen et al. (2019) conducted a meta-analysis on global motion visual processing differences between individuals with ASD and neurotypically developing individuals in behavioural paradigms. They included 48 studies – 28 looked at coherent movement processing from random dot kinematograms and 20 looked at biological motion detection or discrimination of BM from other types of motion (i.e., scrambled). Global motion processing in their context refers to being able to combine several moving stimuli into a coherent shape (i.e., PLDs) or to perceive a coherent direction of the motion of dots despite the existence of unrelated distractor noise. Van der Hallen et al. (2019) found overall differences between ASD and NT individuals in global motion processing but did not find a specific effect for biological motion, rather an effect that indicated a general decreased performance in detecting or recognising global motion

patterns in perception paradigms. Whilst Van der Hallen et al. (2019) found no effect of potential moderators on group differences, they suggest that this may have been due to underpowered studies rather than there being no real effect. However, they did not include emotion processing paradigms and only compared PLDs and random dot kinematograms despite there being other forms of biological motion paradigms, such as animated humans and videos of humans. Another attempt at summarising the behavioural findings in the field was done by Federici and colleagues (2020). They focused on characteristics of PLDs, the levels of processing (first-order/direct/instrumental) and the manipulation of low-level perceptual features in PLDs. They partially answer the question of the effect of the utilised paradigm, showing that when inferring intentions/actions/emotions is required in the task and when temporal manipulations are made to the stimuli, the effects are larger. Unfortunately, their meta-analysis did not focus on the characteristics of the autistic individuals, which, as seen above, have also been suggested to introduce variability in the findings. Finally, while Van der Hallen et al.'s (2019) and Ferderici et al.'s (2020) meta-analyses address the need for a summarisation and exploration of the variability in the results in the literature to a certain extent, their meta-analyses do not fully answer the questions about participant characteristics and their role in the existing findings.

To be able to understand what could drive potential behavioural differences, it is important to also review brain imaging literature for potential answers. There have been some previous attempts to summarise this literature. A meta-analysis on the fMRI investigation of ASD, which included studies on social perception in ASD, found differences between the ASD and NT groups in both basic social tasks such as face recognition and biological motion recognition, and in complex social tasks – i.e., emotion recognition (Philip et al., 2012). However, within social perception, face perception was also included which limits the conclusions that can be made for the perception of only human movement. Similarly, a systematic review by Hamilton (2013) tried to summarise the EEG literature on MNN and autism in BM observation, reporting that experiments probing the relationship between MNN and ASD have produced very mixed results. However, Hamilton (2013) does not provide a quantitative summary of the analysis, only a narrative one.

Since there are inconsistencies in previous findings, behavioural, eye-tracking and brain imaging evidence will be reviewed to identify whether there is substantial evidence for decreased measures of performance in perceiving and understanding BM in individuals on the autism spectrum. We choose to focus solely on biological motion perception as body movement presents qualitatively and perceptually different information from faces and eye-gaze (de Gelder, 2009). Moreover, we want to minimise any inflation or deflation of the effect size of the difference between the two groups, which could be caused by the inclusion of faces and eye-gaze information, which in turn could limit the scope of interpretation. We include studies which have used videos of real humans performing movements, cartoons, which represent humans or human body parts (i.e., hands) (collectively termed full-light displays), and PLDs as described above. The inclusion of both behavioural and physiological measures will allow us to develop a comprehensive understanding of the differences between ASD and NT individuals. Where enough data were available (only in behavioural studies), we also investigate the effects of different contributing factors such as the age, sex and IQ of the participants, the quality of the studies and the effect different paradigms might have on the size and direction of the effect sizes.

3.3 Methods

3.3.1 Protocol

Before commencing this meta-analysis an informal protocol was agreed by all authors based on PRISMA guidelines (Moher et al., 2015). Following these guidelines, the protocol includes details about the methodology and the steps taken to collect and analyse the data, which were agreed prior to commencing this meta-analysis. Through discussions throughout the meta-analytic process and as problems arose, small changes were agreed upon by all authors, such as the exact analysis software, publication bias measures, age categories, etc. The changes are indicated within the protocol. The protocol is available upon request.

3.3.2 Study selection

In order to identify eligible studies, we conducted a systematic literature search. The computerised search involved using the following electronic databases: Dissertations & Theses A&I (ProQuest), Dissertation & Theses: UK & Ireland (ProQuest), Web of Science, PsycINFO (EBSCOhost) and MEDLINE (OVID). The following search terms were used 'autis*', 'biological motion', 'human motion', 'asd', 'asperger*', 'childhood schizophrenia', 'kanner*', 'pervasive development* disorder*', 'PDD-NOS', 'PDD*', 'PLD*','point-light display*', "action observation*", "action observation network*", 'AON'. The asterisk represents truncation, allowing the search to find items containing different endings of the term. Dissertations and Theses databases were searched in order to identify unpublished experiments in an attempt to minimise bias. The search was limited to results in English. Appendix A shows the search strategies used and the number of results the search returned. The search included a wide time span as no lower time criterion was imposed on the search engines allowing us to access the first available records. Results included records up to and including the first week of November 2017. A second search was done in May 2019 for any additional records, due to the substantial time that had passed from the initial search.

The following exclusion/inclusion criteria were then used when screening the remaining records' abstracts and full text:

1. Published before week one of November 2017(Search 1) and May 2019 (Search 2);

2. Published primary empirical articles and theses with non-published results – excluding review articles, opinion pieces, correspondences, case studies, and meta-analyses;

3. Participants in the sample must have an ASD diagnosis;

4. Diagnosis must be confirmed through ADOS, ADI-R or a clinician;

4.1. Added during review process: additional diagnostic measures such as the 3-Di, DISCO; those that are specific to Asperger's disorder, for example the Gilliam Asperger Disorder Scale (GADS, as cited in Price et al., 2012), the Asperger Syndrome (and high functioning autism) Diagnostic Interview (ASDI as cited in Price et al., 2012), and the high-functioning Autism Spectrum Screening Questionnaire (ASSQ as cited in Price et al., 2012) were also accepted as confirmation of ASD diagnosis. Additionally, the Chinese/Japanese equivalents of tests were accepted as in Wang et al. (2015) and Fujisawa et al. (2014).

5. Study must contain fMRI, EEG, eye-tracking and/or behavioural designs;

6. An ASD and NT control group must be present and compared;

7. Although human biological motion includes face motion and eye-gaze, only papers involving human body movement were included to provide a more focused review. These include full-light displays and PLDs;

8. When stimuli that aim to minimize the availability of structural cues (e.g., PLDs) were used, the stimuli must represent human form with a minimum of two points for PLDs;

9. Studies that used videos of people or cartoons where the face was not obstructed were not included as faces could confound with the participants' performance;

10. Papers that focus on imitation of biological motion were not included;

11. If papers focusing on imitation included a separate analysis of BM observation, solely the BM observation was included where possible;

12. Similarly, if paradigms included additional stimuli, but performance on the BM paradigm was analysed and could be extracted separately from the other stimuli, only that analysis was included;

13. Only papers that included t-statistics, descriptive statistics and/or effects sizes were included. Data requests were made to authors, where eligible papers did not include the necessary data.

Two reviewers independently screened the titles, abstracts and full texts against the eligibility criteria. Disagreements were discussed and resolved by the two reviewers or by consultation with the third author. The final decisions on inclusion/ exclusion of the studies were compared between the two reviewers. Cohen's Kappa at the first search was calculated which equated to 64.07%. However, since Cohen's Kappa is sensitive to distribution inequality (Byrt et al., 1993) and ~92% of the records were classified as false positives, the prevalence index (0.816) and the prevalence-adjusted and bias-adjusted kappa (PABAK) of inter-rater reliability were calculated (PABAK = 87.98% inter-rater reliability, absolute agreement = 93.99%). To minimise effort at the second search, inclusion/exclusion was compared at abstract level and then at full-text level (Abstract level: Kappa = 70.72%, PABAK = 80.33%; Full-text: Kappa = 69.57%, PABAK = 71.43%)

The references of included records were screened by hand, split between the two reviewers. Five further records were identified.

3.3.3 Coding and data extraction

Coding of the studies was split between the first and second author. The studies were not double coded, however, the studies coded by the second author were double-checked by the first author. Papers were coded and data was extracted for the following variables:

- Sample size for each group;
- Age: Mean and Standard deviation were extracted for both the NT and ASD groups and each group was post-hoc classified into one of three age groups – children (≤ 13), adolescents (> 13 and ≤ 19) and adult (> 19);
- Full-Scale IQ: Mean and standard deviation were extracted for both the NT and ASD groups;

- Non-verbal IQ: Mean and standard deviation were extracted for both the NT and ASD groups;

- Sex ratio: the sex ratio for each group was extracted and transformed into the proportion of females present in the sample;

- Paradigm: the type of paradigm used was extracted and categorised as: 1 – Detection of biological motion in noise or in comparison to another stimulus (usually upside down or scrambled PLD) (Annaz et al., 2010; Price et al., 2012; Saygin et al., 2010); 2 – Action and subjective states categorisation or recognition (Nackaerts et al., 2012; Van Boxtel et al., 2016; L.-H. Wang et al., 2015); 3 – Emotional states categorisation (A. P. Atkinson, 2009; Fridenson-Hayo et al., 2016; Philip et al., 2010); 4 – Passive viewing (only relevant in fMRI, EEG and eye-tracking). What category each study falls in can be seen in Table 3 and Table 4. Although we initially attempted to separate detection in noise from recognition in comparison to other stimuli, the authors later decided that both tasks would require a similar process of integrating low level information into a coherent human form to perform the task. Thus, to create balanced categories and conceptually cohesive categories the two categories were combined.

- Type of stimulus: the stimuli were grouped into two categories: 1 – PLDs; 2 – Full-light displays – videos of real people or animations.

- Data on performance in the sense of descriptive statistics, *t*-values or effect sizes (*d*), were extracted from each paper. Effect sizes for thresholds, accuracy, sensitivity indices, error rates and reaction times were recorded from the behavioural studies. The areas of activation with contrasts of ASD>NT or NT>ASD were recorded from the fMRI studies and fixations, or proportion of fixations were collected from the eye-tracking experiments. Eye-tracking studies included preferential looking paradigms in which percentage fixations were recorded as an indication of preference for one display, i.e., BM, over another, i.e., inverted BM. Differences in EEG-recorded activation between the NT and ASD groups

were extracted from the EEG experiments, along with the specific frequencies and electrodes used.

Additionally, the following variables were extracted to allow for a complete account of the included studies and quality assessment:

- Diagnosis confirmation criteria;
- Type and number per diagnosis category (where available);
- Additional diagnoses reported;
- Verbal IQ and other cognitive abilities that were not measured by a complete IQ assessment;
- Length of presented stimulus;
- Quality assessment

Risk of bias for behavioural, eye-tracking and EEG studies was assessed by two independent reviewers using the Standard Quality Assessment (SQA) criteria for evaluating primary research papers from various fields for quantitative studies (Kmet et al., 2004). The checklist contains 14 items. Items 5 (If interventional and random allocation was possible, was it described?), 6 (If interventional and blinding of investigators was possible, was it reported?), 7 (If interventional and blinding of subjects was possible, was it reported?) were not used as they refer to the use of interventions which are not applicable for the studies reviewed here. Each of the remaining 11 items can receive 2 points if the assessed study fulfils the criteria; 1 point if it partially fulfils the criteria and 0 points if it does not fulfil the criteria at all. A summary score was calculated for each paper by adding the total score and dividing it by the total possible score. The total score after excluding the previously mentioned three items is calculated with Equation 1. One study (Karuppali, 2018) provided only descriptive information of results (no inferential statistics) and was judged on fewer items (Q1-4, Q8-9, Q13-14).

$$28 - (3[excluded\ items] * 2) = 22 \qquad (1)$$

Eight studies were chosen at random to pilot the quality assessment. Disagreements were discussed and all papers were re-evaluated. An initial comparison was then done between the reviewers' scores. It was found that most disagreements were on item 12 ('*Controlled for confounding?*'). This item was discussed, and the papers were re-evaluated for that item. Disagreements of more than 3 points difference were further discussed on an item-by-item basis. Final comparison of all papers resulted in 18 papers upon which the reviewers completely agreed on the total score. There was no more than a two-point absolute difference between the reviewers' scores for the remaining papers. Thus, the scores for these papers were averaged across both reviewers. Differences between the two reviewers were mostly in the assignment of full or partial points for the items, which was also evident in the original piloting of the scales during its development (Kmet et al., 2004). Overall, the disagreement between the reviewers in the quality score given to each study was quite low with small variability – 0.038(SD = 0.035, min-max [0-0.091]). In total 47 papers were evaluated. The overall SQA score given to all papers was medium/high – 0.792(SD = 0.065, min-max [0.636-0.955]).

We were unable to locate a standardised quality assessment measure that would allow us to assess the quality of fMRI papers. Thus, the assessment was done using relevant criteria from the SQA. Specifically, questions related to the analysis and results were excluded and the fMRI methodology was assessed for robustness. This was done collaboratively by the authors.

For the fMRI studies, which included an analysis of behavioural performance, the fMRI part of the analysis was disregarded initially, and the rest was assessed using the standard SQA procedure described above. This was done to provide a comparable score across the studies that incorporated behavioural performance and to allow for the inclusion of the quality measures as a predictor variable in the analysis. Afterwards, their fMRI protocols and analyses procedures were assessed for methodological robustness by the third and first author. The originally agreed upon score from the SQA was added to the score given for the methodological robustness and a new average quality score was calculated. For the fMRI papers that did not contain a behavioural paradigm, we used the relevant questions from the SQA (Q1-Q4, Q9 & Q12 – Q14). Additionally, their

protocols and analyses procedure were assessed for robustness. These scores were added and a composite score was given. Thus, it is important to underline that the quality scores for the fMRI papers are not directly comparable with the rest of the papers. The quality assessment scores for each study are presented in Table 3 and Table 4.

Additionally, in order to evaluate the quality of the evidence included, we have further conducted a weight of evidence analyses (Gough, 2007). The majority of shortcomings that were identified came from a non-randomised procedure or not including all sample characteristics. Details of this analysis is shown in Appendix B. It indicates that despite their shortcomings, the included studies provide good quality and relevant evidence in support of our conclusions.

3.3.4 Statistical analysis

The following analysis procedure was applied to the behavioural, eye-tracking and EEG experiments. For each included paper, the descriptive statistics, *t*-values or Cohen's *d* were used to calculate Hedges' *g* as the common representation of effect size for all studies. All the calculations and transformations were done by firstly calculating Cohen's *d* and its variance. A correction for small sample size was applied to get the unbiased estimate of Hedges' *g*. The variance of *g* was estimated based on the sample sizes of each study. All the calculations were done using the R package *compute.es* (Del Re, 2013) in R(v3.4.1) (R Core Team, 2019) and RStudio (v.1.1.453) (RStudio Team, 2016). A precision index was calculated for each study as the inverse of the variance (1/variance). Positive Hedges' *g* corresponded to higher scores (better performance) in NT, when compared to ASD. Five top outlier outcomes were identified using a boxplot. An analysis of the initial model with and without the outliers showed that without the outliers, the variance between the studies reduced by a factor of 1.3 and the residual estimates reduced by a factor of five. Thus, all statistical analyses within this paper report the results without the outliers.

Six studies provided RT data. Since a previous meta-analysis (Van der Hallen et al., 2015) showed that RT outcomes tap into different processes in comparison to the rest of the extracted outcomes they were analysed separately from the

rest of the behavioural outcomes. Two top and one bottom outlier were identified using a boxplot. As above, the variance between the studies reduced without the outliers, and the residual estimate reduced by a factor of 3.6. Thus, all statistical analyses report the results without the outliers.

Since papers rarely report only one outcome and/or have only one experiment from which an effect size can be extracted, the traditional (two-level) meta-analysis is not appropriate due to the dependencies that come from using the same subjects or having the same researchers conduct the study (Cheung, 2014; Van den Noortgate et al., 2013, 2015). Therefore, the analysis was extended to a three-level meta-analysis, which takes into account the variance due to the variation of the effect sizes included; the variance that occurs within the same study and the variance that occurs between the studies (Van den Noortgate et al., 2013). Therefore, the three-level analysis estimates these three variance elements. The error only linear model with no moderators as given by Cheung (2014) is shown in Equation 2:

$$g_{jk} = \alpha_0 + u_k + u_{jk} + e_{jk} \quad (2)$$

Where g_{jk} is the effect size for outcome *j* from study *k* and is represented by Hedges' *g*; α_0 is the grand mean of all effect sizes across studies; u_k represents the deviation of the average effect in study *k* from the grand mean; u_{jk} is the deviation of effect *j* in study *k* from the average effect of study *k* and finally e_{jk} is the residual variation not explained by the previously defined variances (Cheung, 2014). This random-effects model is then extended by including moderators. A series of meta-analyses were conducted to investigate the effect of one or a combination of more than one of the following covariates: age, sex ratio, full-scale intelligence quotient (FSIQ) and non-verbal intelligence quotient (NVIQ) for each group, as well as the paradigm and the stimuli. When moderators are added to the analysis, there are two sets of effect sizes that need to be kept in mind. The first set of effect sizes is the difference between ASD and NT at that level of the moderator (or combination of moderators). These are presented in Table 6 and Table 7. The second set of effect sizes are the ones which represent the size of the difference between the different levels. For example, a positive effect size, will indicate that at the first level of

the moderator, the difference between ASD and NT is larger than at the second level. Negative effect sizes here represent that there is a larger effect at the second/third/etc. level than at the previous level.

The parameter estimation was done using maximum likelihood, implemented in the mixed procedure in the statistical package SAS (release 9.04.01, *SAS OnDemand for Academics: SAS Studio*, n.d.). Due to the imbalance of studies when the predictor variables were added, the Satterthwaite method was used to calculate the denominator degrees of freedom (Schaalje et al., 2001). Additionally, to investigate the effects at each level of the categorical variables, a least square means procedure was applied.

To assess heterogeneity, the I^2 statistic (Higgins & Thompson, 2002) was calculated. Since we are using a three-level analysis and potential heterogeneity can occur at the second or the third level, we used the modified formulas provided by Cheung (2014). The I^2 statistic was calculated only for the initial model, the model with the paradigm as a moderator and the model that included both paradigm and age as moderators. This was done because these three models contained the same studies and thus the effect of the moderators on the heterogeneity could be compared. The calculations for level 2 $I^2_{(2)}$ and level 3 $I^2_{(3)}$are shown in Equation 3 below. $I^2_{(2)}$ and $I^2_{(3)}$represent the proportion of variation which can be attributed to the between and within studies respectively.

$$I^2_{(2)} = \frac{\hat{u}^2_{(2)}}{\hat{u}^2_{(2)}+\hat{u}^2_{(3)}+\tilde{v}} \qquad (3)$$

$$I^2_{(3)} = \frac{\hat{u}^2_{(3)}}{\hat{u}^2_{(2)}+\hat{u}^2_{(3)}+\tilde{v}} \qquad (4)$$

Where $\hat{u}^2_{(2)}$ is the between study variance calculated from the model, $\hat{u}^2_{(3)}$ is the within study variance calculated by the model and $\tilde{v}$ is the typical within study variance calculated by Equation 5 as suggested by Higgins and Thompson (2002).

$$\tilde{v} = \frac{\sum w_i(k-1)}{(\sum w_i)^2-\sum {w_i}^2} \qquad (5)$$

Where w is the inverse variance and k is the number of studies.

Publication bias was assessed with Egger Regression (Egger et al., 1997) and the Trim and Fill method (Duval & Tweedie, 2000) using a two-level random effects model. The analysis was performed using a SAS macro created by Rendina-Gobioff and Kromrey (2006).

3.3.5 ALE analysis of fMRI Studies

To analyse the fMRI data, activation likelihood estimation (ALE) in GingerALE v3.0.2 (Eickhoff et al., 2009, 2012; Turkeltaub et al., 2012), was employed. Foci from the between group contrasts, which had reached statistical significance, were first extracted from the studies and converted where necessary into Talairach space using GingerALE. When both whole-brain and region-of-interest analyses were performed, and coordinates were available, the ones from the whole-brain analysis were used. In ALE, the activation foci are shown as a three-dimensional Gaussian probability density function, centred at the specified coordinates. The spatial overlap of these distributions across the different studies and the spatial uncertainty due to inter-subject and inter-experiment variability are then computed. This results in activation maps, which can be seen as summaries of the results of a specified study after considering the spatial uncertainty present. Through the combination of these maps, the convergence of activation patterns across studies can be calculated. This is confined to a grey matter shell and above chance clustering between the studies is calculated as a random-effects factor (Eickhoff et al., 2009). We performed ALE analysis for the NT>ASD contrast only, since only two studies found differences at the ASD>NT contrast (Jack & Morris, 2014; Koldewyn et al., 2011). Only two studies (Alaerts et al., 2013; Grèzes et al., 2009) provided data for emotion detection/identification paradigms, thus this was not analysed separately. Although, our initial intent was to investigate the effects of age, the small number of studies that provided information about the differences between the ASD and the NT group, would not allow for a separate investigation, without introducing spurious results and further complicating the mixed literature in the field. Thus, the readers should keep in mind that the ALE analysis and the output produced contains research from both children/adolescents and adults as well as

emotion and BM detection/observation paradigms. Using the recommended thresholding procedure – cluster defining threshold of 0.001 and cluster-wise family-wise error correction of 0.05 we were not able to identify any significant clusters. An exploratory analysis is reported where we used an uncorrected p-value of 0.001 and maximum cluster size of 200mm^3.

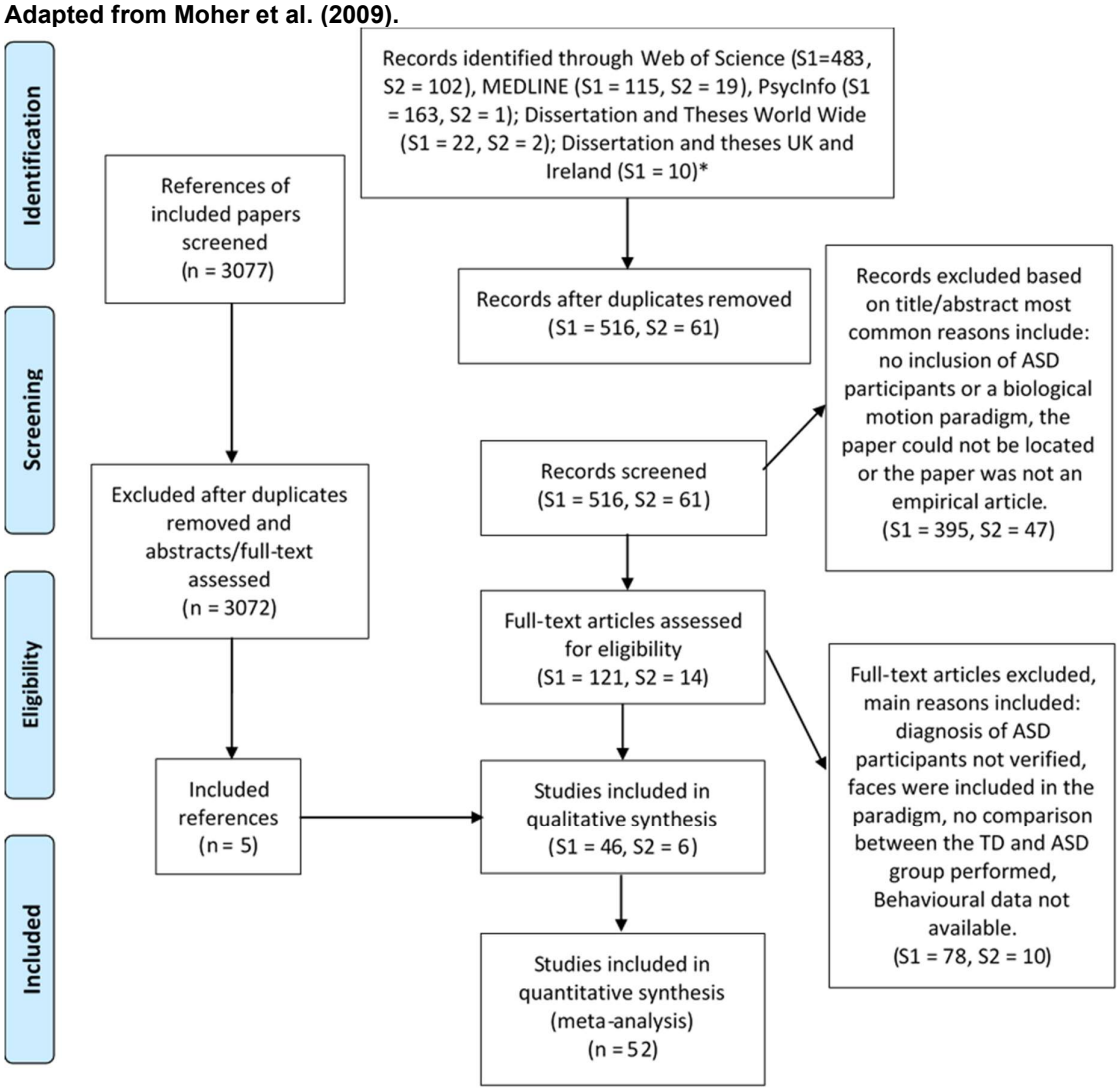
Data used for the analysis is deposited in the ReShare data repository: https://doi.org/10.5255/UKDA-SN-853277.

3.4 Results

The initial (November 2017) study search returned 793 records. The output from all databases was combined and duplicates were removed using two strategies. Initially, R software was used to remove duplicate records that appeared in the same format between the searches. Then, the articles were screened by hand to remove additional duplicates. This resulted in a total of 516 records. At the second search (May 2019), 124 records were identified and Rayyan software was used (Ouzzani et al., 2016). Out of those 45 were identified as duplicates from the previous search and 18 were identified as duplicates between the databases. This resulted in a total of 61 records.

The selection process resulted in a set of 47 papers. Five further records were identified from the references of the included papers. From these 35 contributed to the behavioural studies category, five to the eye-tracking category, five to the EEG category and 11 to the fMRI category. An overview of the inclusion/ exclusion process is shown in the PRISMA flow diagram in Figure 6 below.

Figure 6. PRISMA flow diagram representing the selection/inclusion/exclusion process. Adapted from Moher et al. (2009).



***Note*: The second search did not look into Dissertation and Theses UK & Ireland, as it was covered by Dissertation and Theses Worldwide in the previous search.**

The included studies and their descriptive information can be seen in Table 3 (behavioural, eye-tracking and EEG) and Table 4 (fMRI). The two tables also show the effect sizes for each study, their variance and standard error, their weight of evidence score and their quality assessment score.

Table 3. Summary of behavioural, eye-tracking and EEG studies.

Author (year)	ASD sample					NT sample					Para-digm	Measure	Stimuli	Dura-tion(s)	g	var. g	SE (g)	WoE	SQA
	N	Age	Sex ratio	FSIQ	NVIQ	N	Age	Sex ratio	FSIQ	NVIQ									
Actis-Grosso et al. (2015)	20	22.8	15/5	118.9	/	25	22.3	21/4	/	/	AR	accuracy	PLD (13 strips)	3 (max 5 min)	-0.2	0.09	0.04	8.5	0.71
Alaerts et al. (2015)	12	12/2	13.8	/	100.3	16	13/3	14.2	/	106.7	ER	accuracy	PLD (12 points)	3	0.78	0.15	0.07	8	0.73
	12	12/2	13.8	/	100.3	16	13/3	14.2	/	106.7	ER	RT	PLD (12 points)	3	4.62	0.52	0.14		
	15	15/0	21.7	107.9	105.6	15	15/0	23.3	114.8	109.1	ER	accuracy	PLD (12 points)	3	4.93	0.53	0.13		
	15	15/0	21.7	107.9	105.6	15	15/0	23.3	114.8	109.1	ER	RT	PLD (12 points)	3	-1.53	0.17	0.08		
Alaerts et al. (2017)	15	21.7	15/0	107.9	105.6	15	23.3	15/0	114.8	109.1	D	d'	PLD (12 points)	4	0.66	0.13	0.07	8.5	0.82
Annaz et al. (2010)	23	8.83	/	/	/	34	8.25	/	/	/	D	d'	PLD (13 points)	1	1.15	0.08	0.04	8.5	0.8
	23	8.83	/	/	/	34	8.25	/	/	/	D	thresholds	PLD (13 points)	1	0.89	0.08	0.04		
A. P. Atkinson (2009)	13	30.9	12/1	106.2	105.2	16	26.7	14/2	106.6	108.4	ER	accuracy	PLD (ns) and FLD	3	1.17	0.16	0.07	8.5	0.84
Binnersley (2006)	14	12.917	14/0	/	/	15	/	/	/	/	D	d'	PLD (15 points)	1	0.22	0.13	0.07	8.5	0.8
Blake et al. (2003)	12	/	/	/	/	9	8.417	/	/	/	D	d'	PLD	1	1.13	0.21	0.10	8	0.73
Cook et al. (2009)	16	34.1	14/2	114.8	109	16	33.3	14/2	113	113	D	thresholds	FLD	/	2.82	0.24	0.09	9	0.86
Couture et al. (2010)	36	20.9	29/7	101.3	/	41	22.9	34/7	109.4	/	ER	accuracy	PLD (12 points)	5-20	0.31	0.05	0.03	8	0.84
	36	20.9	29/7	101.3	/	41	22.9	34/7	109.4	/	ER	accuracy	PLD (12 points)	5-20	0.75	0.05	0.03		
Cusack et al. (2015)	15	16.09	15/0	103.1	/	15	15.54	15/0	104.8	/	AR	thresholds	PLD (13 points)	2.5	0.18	0.13	0.07		
	18	16.09	18/0	103.1	/	18	15.54	18/0	104.8	/	AR	thresholds	PLD (13 points)	2.5	-0	0.11	0.06	8	0.82
	18	16.09	18/0	103.1	/	18	15.54	18/0	104.8	/	D	thresholds	PLD (13 points)	1.5	0.3	0.11	0.06		

Author (year)	ASD sample					NT sample					Para-digm	Measure	Stimuli	Dura-tion(s)	g	var. *g*	SE (*g*)	WoE	SQA
	N	Age	Sex ratio	FSIQ	NVIQ	N	Age	Sex ratio	FSIQ	NVIQ									
Edey et al. (2019)	15	16.09	15/00	103.1	/	15	15.54	15/00	104.8	/	D	thresholds	PLD (13 points)	1.5	-0.4	0.13	0.07	8.5	0.91
	15	16.09	15/00	103.1	/	15	15.54	15/00	104.8	/	D	thresholds	PLD (13 points)	2	-0.4	0.13	0.07		
	20	41.1	15/5	115.53	/	17	38.76	14/3	118.24	/	D	d'	FLD	/	-0.05	0.1	0.05		
	22	36.77	18/5	111.18	/	24	31.21	23/1	105.46	/	D	d'	FLD	/	-0.26	0.08	0.04		
Freitag et al. (2008)	15	17.5	13/2	101.2	93.3	15	18.6	13/2	112.1	106.8	D	error rates	PLD	1.5	1.21	0.16	0.07	9	0.84
	15	17.5	13/2	101.2	93.3	15	18.6	13/2	112.1	106.8	D	RT	PLD	1.5	3.14	0.31	0.10		
Frideson-Hayo et al. (2016)	20	7.45	18/2	/	/	22	7.5	19/3	/	/	ER	accuracy	FLD	4-24	0.97	0.1	0.05	8.5	0.95
	20	7.45	18/2	/	/	22	7.5	19/3	/	/	ER	accuracy	FLD	4-24	1.28	0.11	0.05		
	16	8.58	15/1	/	/	18	7.8	13/5	/	/	ER	accuracy	FLD	4-24	1.15	0.13	0.06		
	16	8.58	15/1	/	/	18	7.8	13/5	/	/	ER	accuracy	FLD	4-24	0.31	0.11	0.06		
	19	6.97	15/4	/	/	18	7.36	15/3	/	/	ER	accuracy	FLD	4-24	0.41	0.11	0.05		
	19	6.97	15/4	/	/	18	7.36	15/3	/	/	ER	accuracy	FLD	4-24	0.46	0.11	0.05		
Hubert et al. (2007)	19	21.5	17/2	83.3	/	19	24.33	17/2	/	/	ER	accuracy	PLD (5 & 10 points)	5	2.12	0.16	0.06	7.5	0.64
	19	21.5	17/2	83.3	/	19	24.33	17/2	/	/	AR	accuracy	PLD (5 & 10 points)	5	0.85	0.11	0.05		
	19	21.5	17/2	83.3	/	19	24.33	17/2	/	/	AR	accuracy	PLD (5 & 10 points)	5	1.8	0.14	0.06		
Jones, Swettenham et al. (2011)	89	15.5	81/8	82.1	91.4	52	15.5	49/3	88.4	91.8	D	thresholds	PLD (5 & 10 points)	5	0.3	0.03	0.01	8.5	0.77
Karuppali (2018)	4	5.3	2/2	/	/	4	4.625	2/2	/	/	AR	accuracy	PLD (15 points)	500-2000	3.25	1.04	0.36	7.5	0.69
Koldewyn et al. (2010)	30	15.12	28/2	107.8	104.7	32	15.78	30/2	121.3	114.2	D	thresholds	PLD (13 points)	2	0.82	0.07	0.03	8.5	0.82
	30	15.12	28/2	107.8	104.7	32	15.78	30/2	121.3	114.2	D	RT	PLD (13 points)	2	-0.09	0.06	0.03		
Koldewyn et al. (2011)	16	15.4	14/2	110.6	106.7	16	15.6	14/2	118.6	112.6	D	thresholds (75%)	PLD (13 points)	2	1.01	0.13	0.06	8.5	0.86
Krakowski (2014)	35	10.48	31/4	105.4	108.3	46	11.22	24/22	114	108.2	D	d'	PLD	/	0.6	0.05	0.02	8.5	0.84

Author (year)	ASD sample					NT sample					Para-digm	Measure	Stimuli	Dura-tion(s)	g	var. *g*	SE (*g*)	WoE	SQA
	N	Age	Sex ratio	FSIQ	NVIQ	N	Age	Sex ratio	FSIQ	NVIQ									
Kroger et al. (2014)	17	11.9	17/0	/	/	21	11.63	21/0	/	/	D	accuracy	PLD (15 points)	1	-0.2	0.1	0.05	9	0.86
	17	11.9	17/0	/	/	21	11.63	21/0	/	/	D	RT	PLD (15 points)	1	0.21	0.1	0.05		
Kruger et al. (2017)	16	34.7	12/4	/	/	16	35	12/4	/	/	AR	accuracy	PLD (13 points)	4	-0.1	0.12	0.06	8	0.73
McKay et al. (2012)	10	28.6	10/0	125	/	10	27.9	/	124.8	/	AR	thresholds (50%)	PLD (15 points)	1	0.34	0.19	0.10	8.5	0.73
	10	28.6	10/0	125	/	10	27.9	/	124.8	/	AR	thresholds (75%)	PLD (15 points)	1	0.8	0.2	0.10		
Morrison et al. (2019)	106	24.28	95/11	108.9	/	95	24.17	84/11	116.28	/	D	d'	PLD	/	0.06	0.02	0.01	8.5	0.91
	106	24.28	92/11	108.9	/	95	24.17	84/11	116.28	/	ER	accuracy	PLD	5-10	0.61	0.02	0.01		
Murphy et al. (2009)	16	25.56	13/3	/	/	16	26.4	13/3	/	/	D	d'	PLD (11 points)	~6.5	0.23	0.12	0.06	8.5	0.84
	16	25.56	13/3	/	/	16	26.4	13/3	/	/	D	d'	PLD (11 points)	~6.5	0.32	0.12	0.06		
	16	25.56	13/3	/	/	16	26.4	13/3	/	/	D	d'	PLD (11 points)	~6.5	0.27	0.12	0.06		
	16	25.56	13/3	/	/	16	26.4	13/3	/	/	D	d'	PLD (11 points)	~6.5	0.46	0.12	0.06		
	16	25.56	13/3	/	/	16	26.4	13/3	/	/	D	d'	PLD (11 points)	~6.5	0.29	0.12	0.06		
	16	25.56	13/3	/	/	16	26.4	13/3	/	/	D	error rates	PLD (11 points)	~6.5	0.53	0.12	0.06		
	16	25.56	13/3	/	/	16	26.4	13/3	/	/	D	error rates	PLD (11 points)	~6.5	0.45	0.12	0.06		
	16	25.56	13/3	/	/	16	26.4	13/3	/	/	D	error rates	PLD (11 points)	~6.5	0.37	0.12	0.06		
	16	25.56	13/3	/	/	16	26.4	13/3	/	/	D	error rates	PLD (11 points)	~6.5	0.45	0.12	0.06		
	16	25.56	13/3	/	/	16	26.4	13/3	/	/	D	error rates	PLD (11 points)	~6.5	0.19	0.12	0.06		
	16	25.56	13/3	/	/	16	26.4	13/3	/	/	D	RT	PLD (11 points)	~6.5	0.58	0.12	0.06		

Author (year)	ASD sample					NT sample					Para-digm	Measure	Stimuli	Dura-tion(s)	g	var. *g*	SE (*g*)	WoE	SQA
	N	Age	Sex ratio	FSIQ	NVIQ	N	Age	Sex ratio	FSIQ	NVIQ									
Nackaerts et al. (2012)	16	25.56	13/3	/	/	16	26.4	13/3	/	/	D	RT	PLD (11 points)	~6.5	0.46	0.12	0.06	8	0.8
	16	25.56	13/3	/	/	16	26.4	13/3	/	/	D	RT	PLD (11 points)	~6.5	0.22	0.12	0.06		
	16	25.56	13/3	/	/	16	26.4	13/3	/	/	D	RT	PLD (11 points)	~6.5	0.29	0.12	0.06		
	16	25.56	13/3	/	/	16	26.4	13/3	/	/	D	RT	PLD (11 points)	~6.5	0.27	0.12	0.06		
	12	34.9	7/5	111.5	105.7	12	31.5	7/5	115.5	115.3	ER	accuracy	PLD (12 points)	3	1.71	0.22	0.10		
	12	34.9	7/5	111.5	105.7	12	31.5	7/5	115.5	115.3	D	d'	PLD (12 points)	3	1.28	0.19	0.09		
	12	34.9	7/5	111.5	105.7	12	31.5	7/5	115.5	115.3	D	accuracy	PLD (12 points)	3	1.12	0.18	0.09		
Parron et al. (2008)	23	11.583	20/3	93.7	89	23	12	20/3	/	/	ER	accuracy	PLD (10 points)	~5	5.51	0.41	0.09	8.5	0.73
	23	11.583	20/3	93.7	89	23	12	20/3	/	/	AR	accuracy	PLD (10 points)	~5	1.9	0.12	0.05		
	23	11.583	20/3	93.7	89	23	12	20/3	/	/	AR	accuracy	PLD (10 points)	~5	3.08	0.19	0.06		
Philip et al. (2010)	23	32.5	16/7	101.5	104.4	23	32.4	17/6	111.2	113.4	ER	accuracy	FLD	5 - 10	1.6	0.11	0.05	9	0.84
Price et al. (2012)	14	14.14	14/0	/	57.14	16	14.08	16/0	/	52.63	D	accuracy	PLD	5	1.14	0.15	0.07	8.5	0.8
	14	14.14	14/0	/	57.14	16	14.08	16/0	/	52.63	D	d'	PLD	5	0.82	0.14	0.07		
	14	14.14	14/0	/	57.14	16	14.08	16/0	/	52.63	D	d'	PLD	5	0.64	0.13	0.07		
Saygin et al. (2010)	16	33.75	13/3	112.2	107.2	20	37.75	14/7	113.2	108.6	D	thresholds	PLD (12 points)	0.583	0.08	0.11	0.06	8	0.73
Sotoodeh et al.(2019)	20	11.3	17/3	/	77.3	20	11.4	17/3	/	106	AR	accuracy	PLD (13 points)	2s (3 times)	0.85	0.1	0.05	9	0.82
	20	11.3	17/3	/	77.3	20	11.4	17/3	/	106	AR	RT	PLD (13 points)	2s (3 times)	1.85	0.14	0.06		
Swettenham et al. (2013)	13	9.583	/	/	/	13	8.5	/	/	/	AR	accuracy	PLD (13 points)	/	0.39	0.15	0.08	8.5	0.77
Turi et al. (2017)	19	11.49	16/3	105.9	/	18	11.94	14/4	107.6	/	AR	d'	PLD (23 points)	0.90 ± 0.15	1.64	0.14	0.06	9	0.8

Author (year)	ASD sample					NT sample					Para-digm	Measure	Stimuli	Dura-tion(s)	g	var. *g*	SE (*g*)	WoE	SQA
	N	Age	Sex ratio	FSIQ	NVIQ	N	Age	Sex ratio	FSIQ	NVIQ									
van Boxtel et al. (2016)	19	11.49	16/3	105.9	/	18	11.94	14/4	107.6	/	AR	d'	PLD (23 points)	0.90 ± 0.15	0.44	0.11	0.05	9	0.82
	16	14.04	12/4	101.5	99.94	17	13.32	13/4	112.2	108.6	AR	thresholds	PLD	1	0.27	0.12	0.06		
von der Luhe et al. (2016)	16	41.56	12/4	116.9	/	16	36.19	10/6	115.3	/	D	accuracy	PLD	3.6 - 4.3	0.47	0.12	0.06	8.5	0.89
	16	41.56	12/4	116.9	/	16	36.19	10/6	115.3	/	D	accuracy	PLD	3.6 - 4.3	0.53	0.12	0.06		
Wang et. (2015)	21	5.588	17/4	/	/	21	4.95	16/5	/	/	AR	accuracy	PLD (11 points)	2	1.07	0.11	0.05	8.5	0.8
Eye-tracking																			
Annaz et al. (2012)	17	5.583	/	/	/	17	5.5	/	/	/	PV	% fixations	PLD	6	1.19	0.13	0.06	8.5	0.73
	17	5.583	/	/	/	17	5.5	/	/	/	PV	% fixations	PLD	6	1.75	0.16	0.07		
Burnside et al. (2017)	16	5.22	16/0	/	/	16	3.98	8/8	/	/	PV	% fixations	PLD (13 points)	6	0.29	0.12	0.06	8.5	0.82
Fujioka et al. (2016)	21	27.6	21/0	99.8	96.4	35	25.2	35/0	/	/	D	% fixations	PLD	20	0.32	0.08	0.04	8	0.8
Fujisawa et al. (2014)	15	4.824	12/3	/	/	58	4.008	27/31	/	/	D	% fixations	PLD	20	-0.2	0.08	0.03	8	0.73
Nackaerts et al. (2012)	12	34.9	7/5	111.5	105.7	12	31.5	7/5	115.5	115.3	ER	fixation time	PLD (12 points)	3	3.27	0.38	0.13	8	0.8
	12	34.9	7/5	111.5	105.7	12	31.5	7/5	115.5	115.3	AR	fixation time	PLD (12 points)	3	2.58	0.29	0.11		
EEG																			
Bernier et al. (2007)	14	23.6	0/14	114	107.6	15	26.7	0/15	108.9	105.4	PV	mu-power ratio (8-13HZ)	FLD	2.5 - 20	0.88	0.14	0.07	8.5	0.84
Bernier et al. (2013)	19	6.4	18/1	118.3	112	19	6.9	17/2	95.5	96.3	PV	mu-power ration (8-13HZ)	FLD	6	-0.06	0.11	0.05	9	0.77
Hirai et al.(2014)	12	15.7	11/1	/	/	12	16.5	11/1	/	/	PV	amplitude N100	PLD	0.990	-0.21	0.16	0.08	8.5	0.77
	12	15.7	11/1	/	/	12	16.5	11/1	/	/	PV	amplitude N100	PLD	0.990	0.71	0.17	0.08		

Author (year)	ASD sample					NT sample					Para-digm	Measure	Stimuli	Dura-tion(s)	g	var. *g*	SE (*g*)	WoE	SQA
	N	Age	Sex ratio	FSIQ	NVIQ	N	Age	Sex ratio	FSIQ	NVIQ									
Dumas et al. (2014)	10	33.9	7/3	/	/	30	28.7	14/16	/	/	PV	mu-power ratio (8-13HZ)	FLD	6	3.29	0.26	0.08	8.5	0.73
	10	33.9	7/3	/	/	30	28.7	14/16	/	/	PV	mu-power ratio (8-13HZ)	FLD	6	2.83	0.23	0.08		
	10	33.9	7/3	/	/	30	28.7	14/16	/	/	PV	mu-power ratio(8-13HZ)	FLD	6	2.76	0.22	0.07		
	10	33.9	7/3	/	/	30	28.7	14/16	/	/	PV	mu-power ratio (8-13HZ)	FLD	6	0.05	0.13	0.06		
	10	33.9	7/3	/	/	30	28.7	14/16	/	/	PV	mu-power ratio (8-13HZ)	FLD	6	1.95	0.18	0.07		
	10	33.9	7/3	/	/	30	28.7	14/16	/	/	PV	mu-power ratio (8-13HZ)	FLD	6	-0.33	0.13	0.06		
	10	33.9	7/3	/	/	30	28.7	14/16	/	/	PV	mu-power ratio (8-10HZ)	FLD	6	0.17	0.13	0.06		
	10	33.9	7/3	/	/	30	28.7	14/16	/	/	PV	mu-power ratio (8-10HZ)	FLD	6	0.79	0.14	0.06		
	10	33.9	7/3	/	/	30	28.7	14/16	/	/	PV	mu-power ratio (8-10HZ)	FLD	6	-0.64	0.13	0.06		
	10	33.9	7/3	/	/	30	28.7	14/16	/	/	PV	mu-power ratio (8-10HZ)	FLD	6	-0.6	0.13	0.06		
	10	33.9	7/3	/	/	30	28.7	14/16	/	/	PV	mu-power ratio (8-10HZ)	FLD	6	-0.9	0.14	0.06		
	10	33.9	7/3	/	/	30	28.7	14/16	/	/	PV	mu-power ratio (8-10HZ)	FLD	6	-1.31	0.15	0.06		

Author (year)	ASD sample					NT sample					Para-digm	Measure	Stimuli	Dura-tion(s)	g	var. *g*	SE (*g*)	WoE	SQA
	N	Age	Sex ratio	FSIQ	NVIQ	N	Age	Sex ratio	FSIQ	NVIQ									
	10	33.9	7/3	/	/	30	28.7	14/16	/	/	PV	mu-power ratio (11-13HZ)	FLD	6	4.96	0.44	0.10		
	10	33.9	7/3	/	/	30	28.7	14/16	/	/	PV	mu-power ratio (11-13HZ)	FLD	6	4.92	0.43	0.10		
	10	33.9	7/3	/	/	30	28.7	14/16	/	/	PV	mu-power ratio (11-13HZ)	FLD	6	5.52	0.51	0.11		
	10	33.9	7/3	/	/	30	28.7	14/16	/	/	PV	mu-power ratio (11-13HZ)	FLD	6	3.98	0.33	0.09		
	10	33.9	7/3	/	/	30	28.7	14/16	/	/	PV	mu-power ratio(11-13HZ)	FLD	6	-4.33	0.36	0.09		
	10	33.9	7/3	/	/	30	28.7	14/16	/	/	PV	mu-power ratio (11-13HZ)	FLD	6	1.9	0.17	0.07		
	20	11.158	18/2	103.2	/	19	10.719	14/5	112.7	/	PV	mu-power ratio (8-13Hz)	FLD	80	0.47	0.1	0.05		
Raymaekers et al. (2009)	20	11.158	18/2	103.2	/	19	10.719	14/5	112.7	/	PV	mu-power ratio (8-13Hz)	FLD	80	0.48	0.1	0.05	9	0.8
	20	11.158	18/2	103.2	/	19	10.719	14/5	112.7	/	PV	mu-power ratio(8-13Hz)	FLD	80	0.91	0.11	0.05		

***Note*: N - sample size; FSIQ - full-scale IQ; NVIQ - non-verbal IQ; AR - Action Recognition; D - BM detection; ER - emotion recognition; PV – passive viewing; FLD – full-light display; PLD - point-light display; *d'* –sensitivity index; *g* - Hedges' *g*; var. *g* – estimated variance of *g*; SE(*g*) - estimated standard error of *g*; WoE - weight of evidence; SQA – standard quality assessment score. ▲ - Papers that only reported performance index (accuracy/RT) or in addition to other findings.**

Table 4. Summary of fMRI studies.

Authors (year)	ASD sample					NT sample					Para-digm	Mea-sure	Stimuli	Dura-tion (s)	Contrast		Areas of activation	WoE	SQA
	N	Age	Sex ratio	FSIQ	NVIQ	N	Age	Sex ratio	FSIQ	NVIQ					Task	Groups			
Alaerts et al. (2013)	15	21.7	15/0	107.9	105.6	15	23.3	15/0	114.8	109.1	ER	WB	PLD (12 points)	3s	EM>fix	NT>ASD	L IPL, R MTG-pSTS, L MOG, L MTG-pSTS	8	0.7▲▲
Alaerts et al. (2017)	15	21.7	15/0	107.9	105.6	15	23.3	15/0	114.8	109.1	ER/D	WB	PLD (12 points)	4s	EM>fix	NT>ASD and ASD>NT	None	8.5	0.79▲▲
Bjorndotter et al. (2016)	37	boys: 11.45 girls: 10.73	27/10	boys: 91.33 girls: 98.8	/	38	boys: 11.52 girls: 11.51	25/13	boys: 105.46 girls: 93.85	/	PV	WB	PLD (16 points)	24s	BM>SCR	NT>ASD	R FFG, L FFG, L MTG, L IFG, L Cerebellum	8.5	0.8▲
Freitag et al. (2008)	15	17.5	13/2	101.2	93.3	15	18.6	13/2	112.1	106.8	D	WB	PLD (15 points)	1.5s	BM>SCR	NT>ASD	R Calcarine sulcus, R Parieto-occipital sulcus, R Central sulcus/Postcentral gyrus, R Postcentral gyrus/ Postcentral sulcus, R Postcentral sulcus/ IPL, R IPL, R MTG/STS, R Insula, R ACG, R MedFG, R MFG, L Central sulcus/ postcentral gyrus, L IPL, L FFG, L STG, L Claustrum, L ACG	8.5	0.81▲▲
Grezes et al. (2009)	12	26.6	10/2	102	/	12	21	12/0	119	/	AR/ER	WB	FLD	3s	Dynamic (fear & neutral) vs Static	NT>ASD	R TPJ/STG, R ITG (MT), L ITG, R medSFG, R STG (middle part), R Precentral gyrus, L IFG, R Precaneus, R MFG, R ITG (MT), R FFG/ Cerebellum	9	0.85▲
Jack and Morris (2014)	15	14.2	13/2	110.53	110.33	15	13.8	11/4	112.27	107.6	PV	WB	FLD	7-9s	Hand vs baseline	ASD>NT	R Transverse Temporal Gyrus, L Superior Temporal Gyrus, L Precaneus, R Anterior	9	0.9▲

Authors (year)	ASD sample					NT sample					Para-digm	Mea-sure	Stimuli	Dura-tion (s)	Contrast		Areas of activation	WoE	SQA
	N	Age	Sex ratio	FSIQ	NVIQ	N	Age	Sex ratio	FSIQ	NVIQ					Task	Groups			
Jack et al. (2017)	35	14.49	27/8	97.6		62	17.3	26/36	101.61		PV	Cereb ellum	PLD (15 points)	24s	BM>SCR	NT>ASD and ASD>NT	Cingulate, L Cingulate Gyrus None	8	0.78▲
Kaiser et al. (2010)	25	11.8	20/5	100.2	98.2	NT: 17 US: 20	NT: 10.9 US: 11.3	NT: 12/5 US: 9:11	NT: 114.1 US: 115.8	NT: 110.1 US: 113. 8	PV	WB	PLD (16 points)	~24s	BM>SCR	NT>ASD	L vlPC, vmPC, R PTS, R Amygdala, R FFG, L FFG	8.5	0.75▲
Koldewyn et al. (2011)	16	15.4	14/2	110.6	106.7	16	15.6	14/2	118.6	112.6	D	WB	PLD (13 points)	2s	BM>COH	NT>ASD	R Insula, Bilateral Caudate, Bilateral Pulvinar; R Intraparietal sulcus, R AG, R STS; R IFS, R MFG, R IFG; L Intraparietal Sulcus, L AG, L STS; Anterior Cingulate Sulcus and Gyrus.	8.5	0.84▲▲
																ASD>NT	R ITG/IOG (2 clusters)		
Marsh and Hamilton (2011)	18	33	/	110.22	104.4	19	32.2	/	113.89	113.4	PV	WB	FLD	24s	Hands>sh apes	NT>ASD	L Middle Cingulate extending to Supplementary Motor Area, L Fusiform/Lingual Gyrus	8	0.73▲
Yang et al. (2017)	31	10.86	31/0	98.1	96.65	17	10.92	17/0	104.1	103.7	PV	WB	PLD (16 points)	~24s	BM>SCR	NT>ASD	R AG, R FFG, R Hippocampus, R IOG, R MOG, R IPG, R ITG, R MTG	8	0.9▲

***Note*: N - sample size; FSIQ - full-scale IQ; NVIQ - non-verbal IQ; AR - Action Recognition; D - BM detection; ER - emotion recognition; PV – passive viewing; WB – whole brain analysis; FLD – full-light display; SCR – scrambled BM; COH – coherent dot motion; PLD - point-light display; WoE - weight of evidence; SQA – standard quality assessment score; L – left; R – right; IPL – inferior parietal lobule; AG – angular gyrus; FFG – fusiform gyrus; IOG – inferior occipital gyrus; MOG – middle occipital gyrus; ITG – inferior temporal gyrus; MTG – middle temporal gyrus; (p)STS – (posterior) superior temporal sulcus; (med)SFG – (medial)superior frontal gyrus; IFG – inferior frontal gyrus; MFG – middle frontal gyrus; vlPC – ventrolateral prefrontal cortex; PTS – posterior temporal sulcus; vmPC – ventromedial prefrontal cortex; TPJ – temporo-parietal junction; STG – Superior temporal gyrus; ACG – anterior cingulate gyrus; ▲▲ - Score represents the total score obtained from the behavioural quality assessment plus a score given for the fMRI protocol; ▲ - Score represents the relevant questions from the quality assessment measure + a score for the fMRI protocol.**

This meta-analysis examined 52 papers, which contributed 80 (11 RT) behavioural effect sizes, seven eye-tracking effect sizes, 25 EEG effect sizes and 76 fMRI Foci. The sample size for the behavioural sample included 1742 subjects (ASD: 867, NT: 875). The complete eye-tracking sample included a total sample of 217 participants (ASD: 65, NT: 122). The EEG sample had a total sample of 170 participants (ASD: 75, NT: 95). The fMRI sample had a total sample of 483 participants (ASD: 234, NT: 249). Participant characteristics from all studies (including studies considered outliers in the analyses) are shown in Table 5.

Table 5. Participant characteristics in each type of analysis.

Paradigm (number of studies)	Included Studies	ASD					NT					
		Age (SD)	Proportion of females mean (SD)	FSIQ mean (SD)	NVIQ mean (SD)	N	Age (SD)	Proportion of females mean (SD)	FSIQ mean (SD)	NVIQ mean (SD)	N	N
Behavioural (N=35)	(Actis-Grosso et al., 2015; Alaerts et al., 2015, 2017; Annaz et al., 2010; A. P. Atkinson, 2009; Binnersley, 2006; J. Cook et al., 2009; Couture et al., 2010; Cusack et al., 2015; Edey et al., 2019; Freitag et al., 2008; Fridenson-Hayo et al., 2016; Hubert et al., 2007; Jones, Swettenham, et al., 2011; Karuppali, 2018; Koldewyn et al., 2010, 2011; Krakowski, 2014; Kröger et al., 2014; Krüger et al., 2017; McKay et al., 2012; Morrison et al., 2019; P. Murphy et al., 2009; Nackaerts et al., 2012; Philip et al., 2010; Price et al., 2012; Saygin et al., 2010; Sotoodeh et al., 2019; Swettenham et al., 2013; Turi et al., 2017; Van Boxtel et al., 2016; Von Der Lühe et al., 2016; L.-H. Wang et al., 2015)	19.86 (10.75)	19.15 (27.69)	106.3 (9.76)	98.28 (13.58)	867	19.46 (10.28)	23.38 (23.85)	111.93 (7.42)	105.28 (15.03)	875	1742

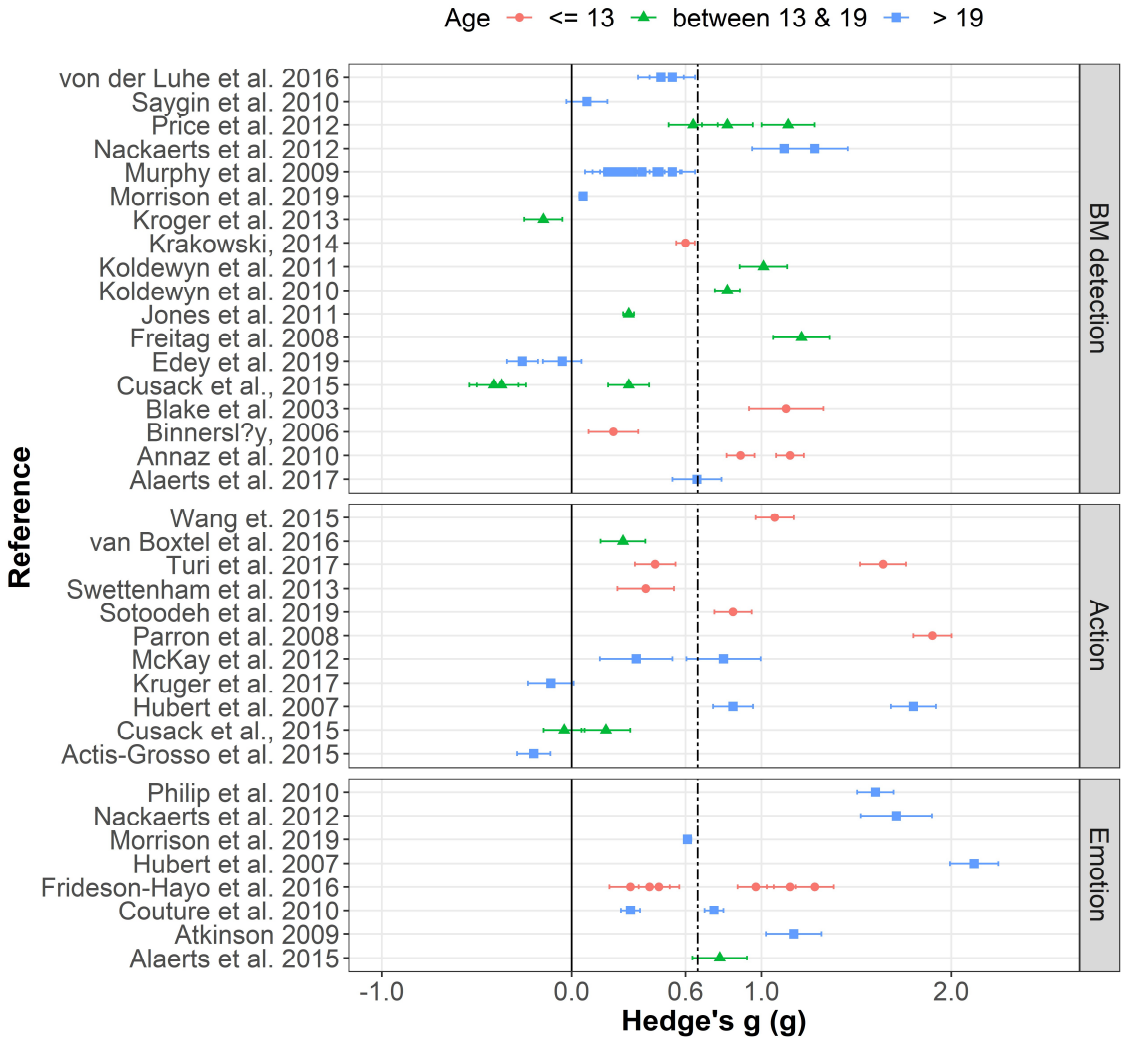
Paradigm (number of studies)	Included Studies	ASD					NT					
		Age (SD)	Proportion of females mean (SD)	FSIQ mean (SD)	NVIQ mean (SD)	N	Age (SD)	Proportion of females mean (SD)	FSIQ mean (SD)	NVIQ mean (SD)	N	N
RT (N = 6)	(Alaerts et al., 2015; Freitag et al., 2008; Koldewyn et al., 2010; Kröger et al., 2014; P. Murphy et al., 2009; Sotoodeh et al., 2019)	16.71 (5.76)	9.51 (8.43)	105.63 (3.84)	96.25 (11.66)	123	17.33 (5.76)	10.3 (8.19)	116.07 (4.73)	108.56 (3.36)	135	258
Eye-tracking (N=5)	(Annaz et al., 2012; Burnside et al., 2017; Fujioka et al., 2016; Fujisawa et al., 2014; Nackaerts et al., 2012)	15.63 (14.5)	15.42 (19.89)	105.65 (8.27)	101.05 (6.58)	81	14.04 (13.28)	36.28 (24.69)	115.5 (0)	115.3 (0)	138	217
EEG (N=5)	(Bernier et al., 2007, 2013; Dumas et al., 2014; Hirai et al., 2014; Raymaekers et al., 2009)	18.15 (10.85)	28.61 (38.57)	111.83 (7.78)	109.8 (3.11)	75	17.90 (9.6)	35.49 (36.45)	105.71 (9.05)	100.85 (6.43)	95	170
fMRI (N=11)	(Alaerts et al., 2013, 2017; Björnsdotter et al., 2016; Freitag et al., 2008; Grèzes et al., 2009; Jack et al., 2017; Jack & Morris, 2014; Kaiser et al., 2010; Koldewyn et al., 2011; Marsh & Hamilton, 2011; Yang, Allen, et al., 2017)	18.03 (7.05)	12.91 (9.99)	103.76 (5.8)	102.6 (5.84)	234	17.54 (6.60)	17.76 (21.99)	111.73 (6.41)	109.57 (3.34)	249	483

3.4.1 Behavioural performance

3.4.1.1 Overall

The random effects three-level analysis of the overall sample revealed a mean estimated effect size of g = 0.6639 [SE = 0.0923, 95% CIs: 0.4759 – 0.8520] t(31.6)=7.2, p < 0.0001, which represents a medium effect (Cohen, 1988). Overall, this suggests that ASD participants were less accurate, less sensitive or produced more errors when asked to detect or interpret biological motion in comparison to NT individuals. The between study variance (u_k= 0.1965 [SE = 0.072], Z = 2.73, p = 0.0032) and the within study variance (u_{jk}= 0.0701 [SE = 0.07], Z = 1, p = 0.1584) show that variance occurred mostly between the studies. The heterogeneity at level 2 is $I^2_{(2)}$ = 0.424, which argues for low to moderate heterogeneity and at the third level $I^2_{(3)}$= 0.0539, which falls under the category of low heterogeneity. The variance component was significant only between studies, indicating that the results varied more between than within studies, which mirrors the heterogeneity measures. It can be seen in Figure 7 that the effect sizes of the studies and their confidence intervals cluster around the estimated effect size from the model, and only a few studies cross the line of no difference. Studies included in this analysis are: (Actis-Grosso et al., 2015; Alaerts et al., 2015, 2017; Annaz et al., 2010; A. P. Atkinson, 2009; Binnersley, 2006; Blake et al., 2003; Couture et al., 2010; Cusack et al., 2015; Edey et al., 2019; Freitag et al., 2008; Fridenson-Hayo et al., 2016; Hubert et al., 2007; Jones, Swettenham, et al., 2011; Koldewyn et al., 2010, 2011; Krakowski, 2014; Kröger et al., 2014; Krüger et al., 2017; McKay et al., 2012; Morrison et al., 2019; P. Murphy et al., 2009; Nackaerts et al., 2012; Parron et al., 2008; Philip et al., 2010; Price et al., 2012; Saygin et al., 2010; Sotoodeh et al., 2019; Swettenham et al., 2013; Turi et al., 2017; Van Boxtel et al., 2016; Von Der Lühe et al., 2016; L.-H. Wang et al., 2015).

Figure 7 Forest plot showing the effect sizes (Hedge's g) from each study and its standard error as the error bars of the points.



***Note*: Different colours/shapes represent the different age categories (red/circle – bellow or equal to 13; green/triangle – between 13 and 19; blue/square – older than 19) and the graph is split by paradigm. Solid line represents no effect; positive effect sizes represent instances where ASD participants performed worse than NT; dot-dashed line represents the effect size extracted from the initial model (g = 0.6639).**

3.4.1.2 Quality

An exploratory meta-analysis was run with the quality given to the studies using the quality assessment tool. However, there did not appear to be an effect of the quality of the studies on the results – $F(1,25.6) = 1.79$, $p = 0.1932$. It has to be pointed out that most studies received quite high scores on the quality assessment measure, which could potentially explain the absence of an effect. However, the inclusion of quality did reduce the variation between the studies (u_k= 0.1754 [SE = 0.0696], $Z = 2.52$, $p = 0.0058$), despite slightly increasing the

within-studies variance (u_{jk}= 0.0753 [SE = 0.0767], Z = 0.98, p = 0.1631). For this reason, quality scores were added as a covariate within the rest of the analyses (Scheepers, 2014). For most cases its inclusion either decreased covariance between the studies or had no qualitative effect. All studies from the overall analysis were included in this analysis.

3.4.1.3 Stimuli

To see whether the type of stimuli – full-light or visually sparse (e.g., PLDs) had an effect on participant's performance, the stimuli type was added as a moderator variable. One paper included both full-light displays, and point light displays and thus was excluded (A. P. Atkinson, 2009). This reduced the number of effect sizes for this meta-analysis only from 64 to 63. The analysis showed that there was no overall effect of the type of stimulus used – F(1,24.9) = 0.91, p = 0.3493. Additionally, the effects for full-light displays and PLDs were both significantly above 0 – g = 0.9055 [SE = 0.3055, 95% CIs: 0.2759 – 1.5351] t(24.7) =2.96, p = 0.0066 and g = 0.5842 [SE = 0.1006, 95% CIs: 0.3778 – 0.7905] t(27) =5.81, p < 0.0001, respectively. Full-light displays showed larger variance, potentially due to a smaller number of studies (N=10).

3.4.1.4 Paradigm

There was an overall effect of the type of paradigm used – F(2,61.5) =8.70, p = 0.0005. There was a significant effect of each paradigm type as shown in Table 6, indicating that participants with ASD performed worse than the NT in all paradigms. More interesting are the pairwise differences in performance between the paradigms. The difference in performance between the detection of coherent BM and action recognition/categorisation was not significant (g = -0.0222 [SE = 0.1646, 95% CI: -0.3511, 0.3067], t(63.8) = -0.13, p = 0.8933). However, there were significant differences between the detection of BM and emotion recognition/categorisation (g = -0.5647 [SE = 0.1373, 95% CIs: -0.8399, -0.2896], t(55.8) = -4.11, p = 0.0001), as well as between action recognition/categorisation and emotion recognition/categorisation (g = -0.5426 [SE = 0.1922, 95% CIs: -0.9268, -0.1583], t(62.4) = -2.82, p = 0.0064). In both situations, ASD participants showed more decreased performance in comparison to NT participants in the emotion recognition/ categorisation paradigms than in

any of the other two. After the paradigm was added as a moderator the variance reduced slightly at the between studies level (u_k = 0.1537) and disappeared at the within study level (u_{jk} = 0). Similarly, the heterogeneity, decreased from the initial model for level 2 and for level 3 ($I^2_{(2)}$ = 0.3319 and $I^2_{(3)}$ = 0). Finally, quality scores did not show a significant effect at this stage $F(1,29)$ =3.48, p = 0.0724. All studies from the overall analysis were included in this analysis.

Table 6. Simple effects for each paradigm.

Paradigm	ES	*g*	SE	Lower CI	Upper CI	*df*	*t*	*p>t*
1	36	0.5041	0.1012	0.2989	0.7093	36.4	4.98	<0.0001*
2	17	0.5274	0.1476	0.2316	0.8233	54.7	3.57	0.0007*
3	14	1.0618	0.1422	0.7773	1.3462	60.1	7.47	<.0001*

***Note*: 1 - detection of BM in noise and recognition in comparison to other stimuli; 2 - Action recognition/categorisation; 3 -Emotion recognition/categorisation. ES - number of effect sizes, g - Hedges' g, SE - standard error, df - degrees of freedom. * - significant at 0.05.**

3.4.1.5 Paradigm and age.

Next, both age and paradigm were included in the analyses and were allowed to interact. A meta-analysis with paradigm and age showed no main effects of paradigm ($F(2, 44.2) = 2.10$, $p = 0.1348$) and no interaction between age and paradigm ($F(2, 34.3) = 1.44$, $p = 0.2426$). However, there was a significant main effect of age ($F(2,29) = 3.35$, $p = 0.0492$). Simple effects of each age group are reported in Table 7. Visual representation of the effect sizes is shown in Figure 7, where the graph is separated by paradigm and the different age groups are colour/shape coded. Note that only one effect was recorded for adolescents in the emotion category.

There were no significant differences in the effect size of the ASD-NT difference between adolescents and adults (g = -0.07848 [SE = 0.2178, 95% CIs: -0.5125, 0.7517], $t(42.4) = -0.36$, $p = 0.7204$). However, there were significant differences in the effect size of the ASD-NT difference between children and adolescents (g = 0.5313 [SE = 0.2523, 95% CIs: 0.01878, 1.0438], $t(34.3) = 2.11$, $p = 0.0426$) and between children and adults (g = 0.4528 [SE = 0.1881, 95% CIs: 0.05998, 0.8457], $t(19.7) = 2.41$, $p = 0.0260$). The effects show that in both cases if the tested participants were children, the effects sizes were larger.

After both age and the paradigm were added as moderators the variance between studies reduced even more, with again no variance being attributed to the third level (u_k = 0.0866 and u_{jk} = 0). Furthermore, the heterogeneity was almost completely accounted for by the moderators ($I^2_{(2)}$ = 0.1363 and $I^2_{(3)}$ = 0).

Additionally, the quality scores showed a significant effect– $F(1,30.2) = 8.17$, p = 0.0076, showing that with the increase of the quality of the study, the smaller the effects were. All studies from the overall analysis were included in this analysis.

Table 7. Simple effects for each age group.

Age	ES	*g*	SE	Lower CI	Upper CI	*df*	*t*	*p>t*
1	17	0.9528	01463	0.6443	1.2614	17	6.51	<0.0001*
2	15	0.4215	0.1963	0.02701	0.8160	49	2.15	0.0368*
3	35	0.5000	0.1089	0.2751	0.7249	23.6	4.59	0.0001*

***Note*: Age: 1 - ≤13, 2 - >13 and ≤19, 3 - >19. ES - number of effect sizes, g - Hedges' g, SE - standard error, df - degrees of freedom. * - significant at 0.05.**

3.4.1.6 Sex

The proportion of females in the samples of both ASD and NT participants was included as moderator variables in two smaller meta-analyses. Since several studies did not report information about sex, only 56 effect sizes from 27 studies were included in these analyses. The proportion of females in the ASD sample had no effect on the results ($F(1, 33.2) = 0.11$, $p = 0.7454$) nor did the proportion of females in the NT sample ($F(1, 29.7) = 0.61$, $p = 0.4402$). Studies included in this analysis are: (Actis-Grosso et al., 2015; Alaerts et al., 2015, 2017; A. P. Atkinson, 2009; Couture et al., 2010; Cusack et al., 2015; Edey et al., 2019; Fridenson-Hayo et al., 2016; Hubert et al., 2007; Jones, Swettenham, et al., 2011; Koldewyn et al., 2010, 2011; Krakowski, 2014; Kröger et al., 2014; Krüger et al., 2017; Morrison et al., 2019; P. Murphy et al., 2009; Nackaerts et al., 2012; Parron et al., 2008; Philip et al., 2010; Price et al., 2012; Saygin et al., 2010; Sotoodeh et al., 2019; Turi et al., 2017; Van Boxtel et al., 2016; Von Der Lühe et al., 2016; L.-H. Wang et al., 2015).

3.4.1.7 Full-Scale IQ

Similar to sex, there were several studies that did not report FSIQ for one or both of the groups. For the ones that did report the FSIQ of both ASD and NT participants, FSIQ was also included as a moderator variable in two smaller meta-analyses. These included 18 studies and 30 effect sizes. There was no effect of FSIQ within the ASD sample ($F(1, 15.9) = 0.02$, $p = 0.8889$) nor was there an effect of FSIQ within the NT sample ($F(1, 30) = 3.98$, $p = 0.0553$). Studies included in this analysis are: (Alaerts et al., 2017; A. P. Atkinson, 2009; Couture et al., 2010; Cusack et al., 2015; Edey et al., 2019; Freitag et al., 2008; Jones, Swettenham, et al., 2011; Koldewyn et al., 2010, 2011; Krakowski, 2014; McKay et al., 2012; Morrison et al., 2019; Nackaerts et al., 2012; Philip et al., 2010; Saygin et al., 2010; Turi et al., 2017; Van Boxtel et al., 2016; Von Der Lühe et al., 2016).

3.4.1.8 Non-Verbal IQ

Only 14 studies and 18 effect sizes included the NVIQ for both the ASD and the NT group. Two smaller meta-analyses were performed using the NVIQ of each group as moderator variables, however there were no significant effects neither for the ASD NVIQ ($F(1,12.1) = 0.15$, $p = 0.7012$) nor for the NT NVIQ ($F(1,11.3) = 0.00$, $p = 0.9921$). Studies included in this analysis are: (Alaerts et al., 2015, 2017; A. P. Atkinson, 2009; Freitag et al., 2008; Jones, Swettenham, et al., 2011; Koldewyn et al., 2010, 2011; Krakowski, 2014; Nackaerts et al., 2012; Philip et al., 2010; Price et al., 2011; Saygin et al., 2010; Sotoodeh et al., 2019; Van Boxtel et al., 2016)

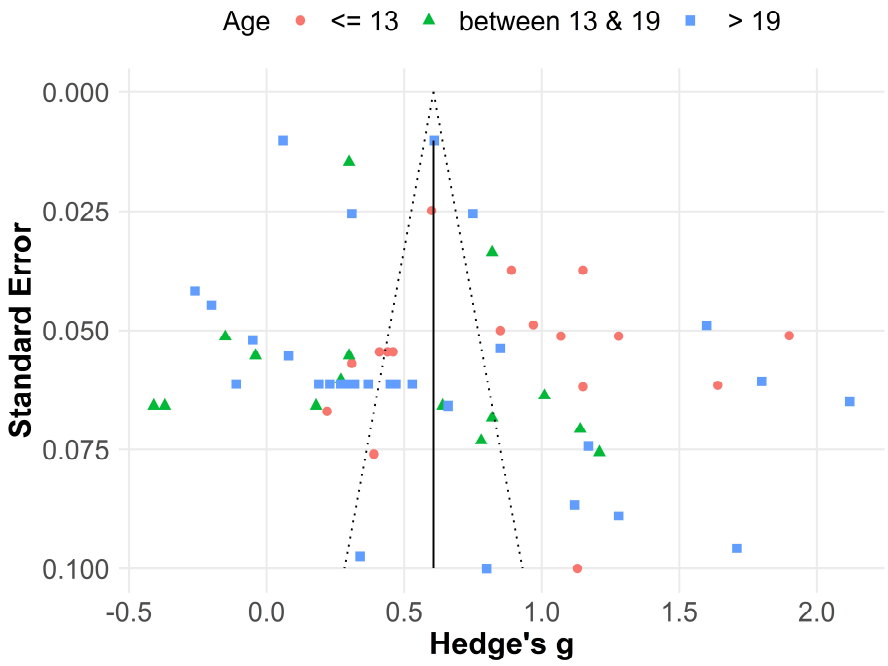
3.4.1.9 Publication Bias

To evaluate the possibility of a publication bias, we plotted the behavioural effect sizes against their standard error with a funnel plot (see Figure 8) (Egger et al., 1997; Sterne & Egger, 2001). As can be seen by their distribution, there is a wide variety of observations with similar standard errors. Specifically, there appears to be a lack of effect sizes with high standard errors and low effect sizes and low standard errors with high effect sizes, which stems from the relatively small to moderate sample sizes in the studies. The inverted funnel shape, which

extends 1.96 standard errors around the overall estimate should include 95% of the studies. However, one of the assumptions for that interpretation is that the true effect is the same in each study (Sterne et al., 2011). It is evident from Figure 3 that 95% of the studies do not fall within the funnel shape. However, we do not make the assumption that the treatment effect is the same in each study. Moreover, we show that the effects vary with age and paradigm. Finally, it is possible that additional variability is added due to the heterogeneous nature of the ASD population.

Besides visual inspection of the funnel plot, the Egger Regression method (Egger et al., 1997) was used to assess the possibility of bias using a random effects model. Egger's regression detected a risk of publication bias – $t = 2.5806$, $p = 0.0122$. Specifically, there is slight asymmetry in the lower end of the funnel plot, where larger standard errors produced larger effect sizes. For this reason, the Trim and Fill method from Duval and Tweedie (2000) was used. Using a standard random effects model, the analysis indicates publication bias in the right tail of the funnel plot, indicating that more studies were published with large effect sizes and large standard errors. This was mirrored by the direction of the effect found in the meta-analysis including the quality assessment scores.

Figure 8. Funnel plot for the behavioural studies.

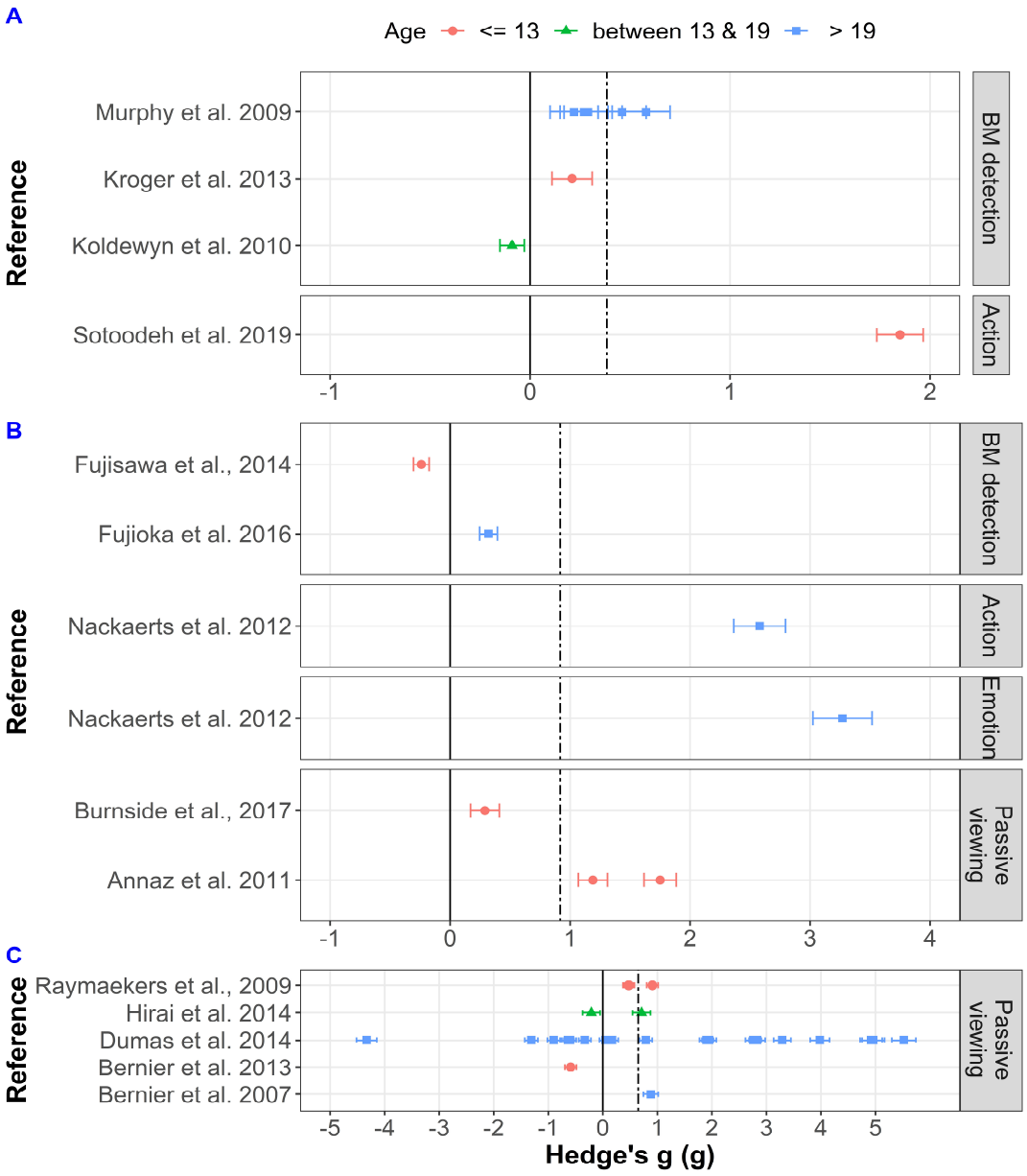


Note: **Displays the effect size – Hedge's g, plotted against the standard error. The vertical line represents the effect size from the overall analysis.**

3.4.2 Reaction time

The random effects three-level analysis of the overall RT sample revealed a mean estimated effect size g = 0.384 [SE = 0.1828, 95% CIs: -0.0375 – 0.8055] $t(8)$=2.1, p = 0.0689, which represents a small effect (Cohen, 1988). Overall, this suggests that ASD participants showed non-significantly slower RT in the BM paradigms in comparison to NT individuals. There was no between study variance (u_k= 0) or within study variance (u_{jk}= 0), thus heterogeneity was not calculated. With the removal of outliers, there were only 8 effect sizes left, and further moderation analyses were not run (Higgins & Green, 2011). Figure 9A shows the distribution of effect sizes for the reaction time paradigms. Studies included in this analysis are:(Koldewyn et al., 2010; Kröger et al., 2014; P. Murphy et al., 2009; Sotoodeh et al., 2019).

Figure 9. Forest plots showing the effect sizes (Hedge's g) from each study and its standard error as the error bars of the points.



Note: **Different colours/shapes represent the different age categories (red/circle – bellow or equal to 13; green/triangle – between 13 and 19; blue/square – older than 19) and the graph is split by paradigm. Solid line represents no effect; positive effect sizes represent instances where ASD participants performed worse than NT; dot-dashed line represents the effect sizes extracted from the initial model. A – Reaction time data (*g* = 0.384), B – Eye-tracking data (*g* = 0.917), C – EEG data (*g* = 0.642).**

3.4.3 Eye-Tracking

As there were only five papers that provided enough information to extract data about effect sizes in eye-tracking experiments, a meta-regression with moderators was not conducted. The five studies contributed a total of seven effect sizes. The overall analysis revealed a mean estimated effect size *g* =

0.9172 [SE = 0.4865, 95% CIs: -0.3552, 2.1896], $t(4.73)=1.89$, $p = 0.1214$, which represents a large effect, but non-significant(Cohen, 1988). Overall, this means, that ASD participants showed less preference for biological motion in comparison to NT individuals, however it should be noted that it was not significant, which is predicated by the broad confidence intervals around the estimate. The between study variance ($u_k = 1.0862$ [SE = 0.7841], $Z = 1.39$, $p = 0.083$) and the within study variance ($u_{jk}= 0.0$) showed that variance occurred mainly between studies, which was expected due to the small number of studies. It is important to point out that due to the small number of studies and the large confidence intervals, these results should be taken with caution. Figure 9B shows the distribution of effect sizes for the eye-tracking paradigms. All studies reported in Table 3 under the eye-tracking subheading are included.

3.4.4 EEG

There were 25 effect sizes provided by five studies. The overall effect size revealed by the analysis was not significant – $g = 0.6489$ [SE = 0.3271, 95% CIs: -0.02476, 1.3226], $t(25) =1.98$, $p = 0.0584$. Similar to the eye-tracking results, this showed a medium effect size but due to the small sample size, and the fact that one study contributed 17 of the effect sizes, it is expected that the large confidence intervals would overlap with 0. There was no between or within study variance- $u_k = u_{jk} = 0$. Figure 9C shows the distribution of effect sizes for the EEG paradigms. Due to the variability that is seen in the frequency that is used, an exploratory analysis, which looks at frequency as a contributing factor to the EEG findings, is reported in Appendix C. All studies reported in Table 3 under the EEG subheading are included.

3.4.5 fMRI

The 11 studies that investigated the difference between ASD and NT participants covered emotion recognition and distinguishing between coherent BM PLD and scrambled PLD/fixation baseline or coherently moving dots. Due to the small sample of studies and the fact that 2 studies did not find any significant brain areas, and one study only found difference in the ASD>NT contrast, all studies were analysed together for the NT>ASD contrast. Only Koldewyn et al. (2011) and Jack et al. (2014) found differences where ASD participants showed

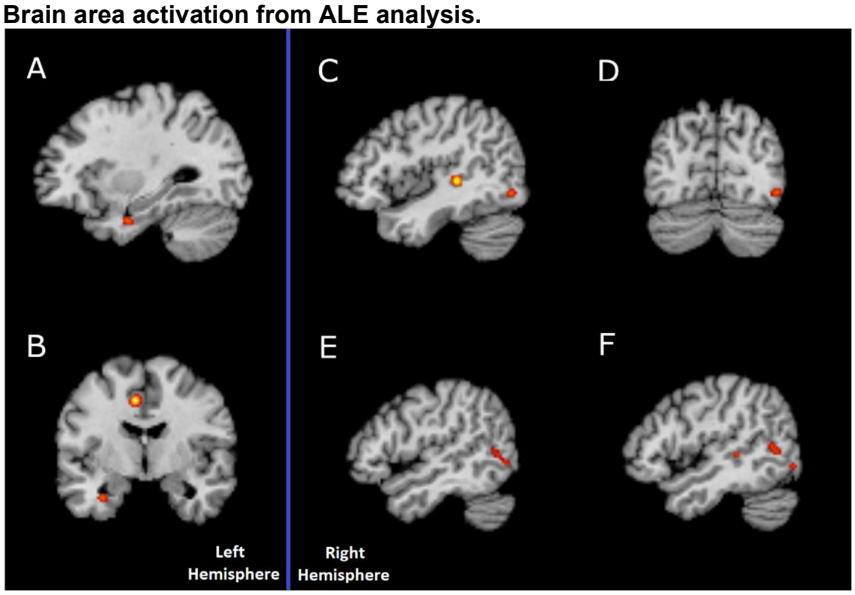
significantly higher activated regions when compared to NT. Since these were the only two studies to show this contrast, no further analysis was done for the ASD>NT contrast. This led to the inclusion of eight studies (62 foci). Due to the small number of included studies, we used the uncorrected p-values at a level of 0.001 and a minimum cluster size of 200mm^3. Table 8 and Figure 10 present the results from the NT>ASD comparison. Five clusters were identified where the NT participants showed greater activation than the ASD participants. In the left hemisphere, one cluster peaked at the left uncus, Brodmann area (BA) 20, and one at the middle cingulate gyrus (MCG), BA 24. The remaining regions were in the right hemisphere, where one region peaked at the middle occipital gyrus (MOG) (BA 19), one region at the superior temporal gyrus (STG) (BA 41) and one cluster with two peaks at the middle temporal gyrus (MTG) and the Inferior Temporal Gyrus (BA 41 and 39 respectively). The resulting map overlays were produced on a standardised structural scan using Mango v4.1 (Lancaster & Martinez, n.d.)(rii.uthscsa.edu/mango).

Table 8. Regions with significantly elevated activation likelihood from the ALE analysis.

Comparison	Cluster	Brain region		BA	Volume (mm^3)	Talairach			ALE (10^{-2})	Range						Centred at		
						x	y	z		from			to					
										x	y	z	x	y	z	x	y	z
NT>ASD	#1	R	STG	41	408	44	-32	4	1.69	40	-34	0	48	-28	6	44.1	-31.3	3.1
	#2	R	MTG ITG	39	312	48 50	-60 -68	8 0	1.07	46	-64	4	50	-56	12	49	-62.2	5.1
	#3	R	MOG	19	264	46	-74	-6	1.20	40	-76	-8	48	-70	-4	44.4	-72.7	-6.2
	#4	L	Uncus	20	248	-32	-4	-28	1.21	-36	-8	-30	-30	0	-26	-32.7	-4.3	-28.1
	#5	L	MCG	24	408	-8	-4	46	1.72	-12	-8	42	-4	0	50	-8	-4.5	45.7

***Note*: BA – Brodmann area; STG - Superior Temporal Gyrus; MTG - Middle Temporal Gyrus; ITG – Inferior Temporal Gyrus; MOG – Middle Occipital Gyrus; MCG – Middle Cingulate Gyrus; R – Right; L – Left.**

Figure 10. Brain area activation from ALE analysis.



***Note*: A – Uncus; B – Central Gyrus, C – Superior Temporal Gyrus, D – Middle Occipital Gyrus; E – Inferior Temporal Gyrus; F – Middle Temporal Gyrus**

3.5 Discussion

The aim of this meta-analysis was to investigate whether ASD individuals show differences in their ability to perceive and interpret biological motion when compared to NT individuals. This question has been under discussion for decades and contradicting results have continuously appeared in the literature. Therefore, a quantitative summary of the results was necessary to allow research to move forward in understanding the atypicalities present in ASD. The current study investigated several potential factors that could contribute to the variable and often mixed results in this field. We explored the possibility of different paradigms being a reason for these varied findings and the effect of age, sex and IQ on participants' performance.

This meta-analysis showed that there is a medium effect indicating an overall decreased performance in perceiving and interpreting biological motion for ASD individuals. Specifically, the present findings show that individuals with autism show lower levels of performance when higher order information, such as emotion, is required to be extracted from biological motion. Moreover, age is a

significant contributing factor to the variability of the results, as different age groups show different degrees of performance decrement. Additionally, we did not find a significant effect in reaction time data, suggesting no delays responding to stimuli once recognised. Further, the effect size of the eye-tracking results would argue that autistic individuals do not attend to or orient towards BM. However, the small sample of studies and its variability led to a non-significant estimated effect size, even though the effect size would be constituted as 'large'. This variability is evident in the distribution of the study effect sizes around the average effect size. Thus, the absence of significance in the eye-tracking results may possibly be mainly attributed to the small sample. A similar pattern is seen from the EEG studies. Finally, the five clusters identified in the fMRI ALE analysis to show higher activation for NT than ASD individuals provide evidence for a potential neural basis for the difference in BM perception abilities.

3.5.1 Differences in performance increase with the increase in task complexity

Biological motion can convey various types of information. It can provide simple information about what others around us are doing, or more complex information, for example about the emotional state of others (Blake & Shiffrar, 2007; Pavlova, 2012). All this information is of great importance in social interaction. Although, Koldewyn et al. (2010) argue that individuals with ASD can perceive/detect biological motion, we found a generally decreased performance in the perception of BM in ASD individuals in all paradigms, including simple BM detection. Moreover, there was no difference in performance between BM detection and action recognition. This indicates that although biological motion detection requires simple integration of motion elements, decreased performance at this level already exists, hindering recognition. Furthermore, the effect size of the difference between the NT and ASD individuals was about twice the size when emotion recognition paradigms were employed. Thus, aligned with Koldewyn et al.'s (2010) arguments, there is in fact decreased performance when the extraction of emotion information is required but this would manifest on top of the already existing decreased performance with simple detection of BM. Similar findings were also observed by Federici et al. (2020), where inferring higher order information from PLDs showed larger

effects. This is an expected finding since ASD is defined with difficulties in social interaction and communication. Emotion recognition is a highly social process, making it more cognitively demanding than BM identification which would rely on perceptual decisions. The effect of paradigm in our meta-analysis may be because emotion adds an additional layer of social complexity in comparison to simple BM identification or action recognition, making it more difficult for individuals with ASD to perform on such tasks. This difference between the two groups is true even when simple and complex emotional recognition tasks are used (Bal et al., 2010; Dyck et al., 2001; Fridenson-Hayo et al., 2016; Hudepohl et al., 2015; but see Jones, Pickles, et al., 2011). One suggested underlying mechanism has been the possibility of underlying alexithymia symptoms in autism, which are related to reduced ability for individuals to recognise their own emotion (Bird & Cook, 2013; R. Cook et al., 2013). This by extension could lead to the increased difficulties in recognising emotional states in others. Hence, co-occurring alexithymia in the studied samples could to an extent be underpinning these results.

It is worth noting that we did not find significant effects when reaction time was the measured outcome. Even more, the effect size that we found would be considered small according to Cohen's (1988) characterisations. Although, a recent meta-analysis has shown that global information integration takes time in autism, which is evident in slower reaction times (Van der Hallen et al., 2015), this is not evident in biological motion perception. A possible explanation is that motion introduces an additional factor, which is suggested by reported higher motion thresholds in autism (Annaz et al., 2010; Pellicano et al., 2005). Moreover, biological motion perception has longer spatiotemporal integration windows than simple motion stimuli, which could make it more difficult to detect small differences in reaction time (Neri et al., 1998). Thus, the decreased performance in perceiving biological motion is a combination between motion and the social factor of human movement, which is more evident in interpretation, rather than in time taken for processing.

This finding, that different paradigms introduce varying effect sizes, emphasizes that when the research community is trying to explain differences between NT and ASD individuals, it cannot simply talk about biological motion perception as a whole. Instead, the nuances that different paradigms bring need to be

emphasised. Moreover, the different paradigms are not comparable; instead, they provide different levels of understanding of the abilities of individuals with ASD.

3.5.2 Differences between ASD and NT individuals decrease with age

The developmental course of BM perception in ASD is critically important, especially since so many contradicting results have been found between different age groups (Actis-Grosso et al., 2015; Krüger et al., 2017; McKay et al., 2012; Parron et al., 2008; Turi et al., 2017; L.-H. Wang et al., 2015). Overall, it appears that the size of the difference between the two groups is larger when children are investigated. On the other hand, the effect size when adults were studied did not differ from the effect size when adolescents were studied.

Our findings suggest that ASD individuals tend to catch up with age and that performance within ASD becomes more aligned with the NT population. This in turn corresponds to the general improvement with age observed within NT individuals (Ghanouni et al., 2015). Despite this catch up however, the size of the differences between the two groups was significant at every age category, indicating consistent difference in performance but to a varying degree dependent on age. Thus, whilst NT and ASD tend to both improve in their ability to detect BM, ASD individuals do so at a slower rate. This implies the existence of a developmental delay in the extraction of relevant social information from biological motion. It should be noted that Annaz et al. (2010) also did not find a relationship with age in children with ASD for non-biological motion coherence and form-from-motion paradigms, whereas the effect was present in NT individuals. Thus, it appears that there might be a global delay in motion coherence sensitivity in ASD. Although, Simmons et al. (2009) argue for inconsistency in the literature about motion coherence and ASD, elevated motion coherence thresholds have been found by others (eg. A. P. Atkinson, 2009; Koldewyn et al., 2010). Moreover, Van der Hallen et al.'s (2019) findings suggest specifically that there is an overall decreased performance in global motion perception in individuals with ASD, for both coherent and biological motion.

To sum, the variability in the behavioural findings in the literature can be explained largely by the fact that ASD participants cannot be put together as a single group. As well as talking about the nuances that individual paradigms bring, we need to distinguish between the different age groups. Thus, a study aiming to investigate performance in adults should not look for effects as large as the ones found in children, as they are statistically not comparable.

3.5.3 No effect of sex, FSIQ and NVIQ on performance on BM paradigms

It has been suggested that ASD is expressed differently in males and females and that females could be the source of variability in some of the results related to performance in the ASD literature (Van der Hallen et al., 2015). However, we did not find any significant effects of the proportion of females in either the NT or ASD sample. Furthermore, neither the FSIQ nor the NVIQ of either group revealed a significant effect on the overall performance. Although some studies have argued for (Jones, Swettenham, et al., 2011; Rutherford & Troje, 2012) and against (A. P. Atkinson, 2009; Van Boxtel et al., 2016; Van der Hallen et al., 2019) the effects of IQ, those that find effects usually have lower IQ scores in comparison to the ones that do not find this effect (but see P. Murphy et al., 2009). The mean FSIQ in the current meta-analysis was also higher – with averages in the behavioural, eye-tracking and fMRI designs falling between 103 and 112. Thus, it is possible that any variability that may be explained from an IQ perspective, might not have been captured in this analysis or in studies where the IQs are above 100. Thus, the present findings may not necessarily be transferable to ASD individuals at the lower end of the IQ distribution. However, since research is usually done on individuals of average or above average IQ, this nuance would not be captured unless more research is adapted and done with individuals on the lower side of the IQ distribution.

3.5.4 Brain and behaviour

From a brain imaging perspective, we aimed to investigate both EEG and fMRI. This was driven by the fact that it has been suggested that individuals with ASD utilise different brain networks when observing biological motion(McKay et al., 2012).

EEG studies, which usually rely on mu-suppression as a proxy for the MNN in ASD, argue for an impaired mirror system in autism (Bernier et al., 2007; Oberman et al., 2005, 2013; Raymaekers et al., 2009). Specifically, they have consistently found reduced mu-suppression in central electrodes. Similar findings have been indicated by a meta-analysis conducted by Fox et al. (2016). However, we did not find a significant effect for the difference between ASD and NT individuals. There are two possible explanations for this result. One possibility is that the effect sizes were too small to be considered significantly different from 0. This, however, does not seem to be the case, as there is a good distribution of results on both sides of the no-difference line. The second possibility is that the small sample of studies, did not provide enough data points to allow for a stable estimate to be given. This is especially evident by the lower bound of the 95% CI for the overall effect size, as it stays very slightly below 0. Furthermore, the exploratory analysis, which is reported in Appendix C, showed that depending on the frequency used to perform the analysis, the effect size can differ greatly. Thus, for some conclusion to be made from the EEG studies, a common analysis structure needs to be agreed upon. However, Hamilton (2013) argues that support for a difference from these studies is weak and mixed, which also speaks for the unreliable findings. Moreover, it has been argued that mu suppression findings can be unreliable as they are very much dependent on the baseline that is chosen (Hobson & Bishop, 2016). Although some of the studies identified here used the same paradigm with the same baseline (Oberman et al., 2005, 2008, 2013), this was not the case for all of them (Bernier et al., 2007; Raymaekers et al., 2009), which makes it difficult to compare the findings. Moreover, a general critique has been given on the association of the mu signal with action observation and imitation, showing that mu signal is better associated with the somatosensory components as opposed to the MNN. Thus, a general standard for data analysis, what constitutes as a baseline, and what is it that mu is measuring needs to be established before any conclusions can be drawn.

From an fMRI perspective we investigated the differences in brain activation between ASD and NT in biological motion perception and emotion recognition. It is noteworthy that emotion perception and BM observation paradigms were analysed together, due to the small sample size. Unfortunately, we were unable to identify significant clusters that overlapped between the studies. However,

the exploratory analysis showed that by using a more relaxed threshold, the areas that come up as different between the two groups correspond to the areas that have been identified in the biological motion perception literature.

In short, we found five clusters where NT individuals showed greater activation than ASD individuals: the left uncus, left middle cingulate gyrus, right middle occipital gyrus and one cluster peaking at the right superior and middle temporal gyri. These findings are consistent with literature showing right hemisphere dominance in the processing of biological motion (Downing & Peelen, 2011; Grosbras et al., 2012). Particularly the right ITG and the right middle temporal gyrus (MTG) have been observed to be specifically implicated in the observation of human motion (Downing, 2001; Grosbras et al., 2012; Noble et al., 2014). Additionally, the ITG has been found to be part of the BM processing network of NT in McKay et al.'s (2012) experiment but not in ASD, which corresponds to our findings. Similarly, the MTG is related to the perception of human movement. Peelen and Downing (2007) argue that the MTG is part of the extrastriate body area (EBA) and that its activation during action observation is due to it representing the shape and posture of the body rather than the action. Additionally, Thompson and Baccus (2012) argue that motion and form make independent contributions to the processing of biological motion in the MT areas. Specifically, the MT areas respond a lot more to the motion aspects, and EBA to the representation of human form. However, since these areas overlap (Thompson & Baccus, 2012) and the observed cluster in these results peaked at MTG and ITG, it could be expected that the activation is due to an interplay between the motion and human form information. This collaborative mechanism has previously been suggested by Downing and Peelen (2011). If individuals with ASD have problems perceiving the basic human shape and posture, it is understandable why there appeared to be consistent differences in behavioural performance between ASD and NT individuals in all biological motion paradigms investigated here. Moreover, as mentioned earlier, with the increased motion thresholds found within individuals with ASD (Pellicano et al., 2005) it could be expected that impairments would come from both motion and human form detection.

Interestingly, the superior temporal sulcus (STS) is a region that has been implied to be important in biological motion perception (Grosbras et al., 2012;

Pavlova, 2012); however, we did not find higher STS activation in NT in comparison to ASD. Nevertheless, we did find the superior temporal gyrus (STG) to have higher activation in NT. Previous findings (Grosbras et al., 2012; Pavlova, 2012; Pelphrey et al., 2011) have argued that the STS is involved in social perception, namely it integrates the social context with the actor's actions. Nevertheless, McKay et al. (2012) also did not find the STS to be involved in simple biological motion perception. Since their paradigm is similar to the paradigms used in the papers, which dominated in the present analysis, it fits that we also did not find clear evidence for STS activation. However, the proximity of the STG to the STS suggests that there might be some potential overlap which could be driven by the inclusion of the emotion related BM paradigms in the analysis. In fact, the STG has been found to show activation when observing emotional biological motion and in biological motion perception paradigms in general (Grosbras et al., 2012; Han et al., 2013; Peelen et al., 2007).

Despite both the low number of studies which were included in the ALE analysis and the exploratory nature of the results, the brain areas found were consistent with BM processing literature. Moreover, differences in these brain areas can and do show differences in behaviour. This finding emphasises the connection between brain differences and behavioural performance. However, due to the small number of studies and the fact that a more constrained threshold did not show any significant values, some caution needs to be taken when interpreting these results.

3.5.5 Methodological limitations

The quality of a meta-analysis is only as high as the quality of the studies that it includes. The studies that we included received a relatively high score on our quality assessment measure with little variance between the studies. The major methodological issues of the included studies were the small sample sizes and the fact that on several occasions there were no corrections for multiple comparisons. However, the correction for multiple comparisons should not have affected our results as we used the descriptive or test statistics, rather than the p values. Nevertheless, it was evident in the behavioural analysis that the quality of the studies played a significant role in reducing variability and

allowing for better interpretability of the statistical results. This indicates that small changes in the quality of a study contributed enough to influence the results. Specifically, it appeared that the higher the quality of a study, the smaller the effect size was indicating that better controlled studies produced smaller effect sizes. The same finding was observed by the publication bias analysis, which showed that studies with smaller standard errors produced smaller effect sizes. This on its own, is an important discovery about the control that is used when developing a study paradigm. It is possible that with a better controlled study, larger amounts of variability are controlled, reducing any additional external effects. Thus, future autism researchers should aim to provide even more methodologically sound results, to allow them to distinguish between external heterogeneity and within-ASD heterogeneity.

Additionally, in our criteria we aimed to include studies that utilised either the gold standard (i.e. ADOS plus ADI; see Simmons et al., 2009) or expert clinical opinion when confirming the ASD diagnosis of their participants. However, during the selection process we realised that a number of studies did not employ the gold standard and rather used various diagnostic measures. For that reason, we expanded our inclusion criteria to include at least some form of diagnosis confirmation. Worryingly, one of the reasons that studies were not included in the present analysis was that the diagnosis was not confirmed by any means, let alone by using the gold standard. However, the concept of a gold standard is a matter of debate (Matson & Neal, 2009) and it has been noted that the scales do not always capture individuals that have been diagnosed with Asperger's syndrome (Price et al., 2012). Thus, how ASD participants ought to be identified in future studies needs to be explored.

Furthermore, even though it is argued that a quantitative summary on two effect sizes is better than simple counts of positive vs negative effects (Valentine et al., 2010), statistical analysis, and the confidence one can give to it, is proportionally dependent on its sample size. Although the three-level model has allowed us to utilise more than one effect size per study, thus increasing the number of cases included, the resulting sample is still small, especially for some of the categories of analysis. This is mainly true for the EEG analysis, where one study provided most of the effect sizes. Thus, when interpreting the results from this meta-analysis, the number of studies in each

part needs to be considered. Furthermore, the number of effect sizes that we were able to include in some of the analyses (eye-tracking, RT, EEG and fMRI) did not allow us to investigate important factors such as paradigm and age. This unfortunately limits our ability to interpret the effect of those factors. Nevertheless, if we look at the behavioural results, then we can conjecture, that these factors will be important and will also need to be considered, when new paradigm designs are considered, or when interpreting the overall weight of the effects found in the literature.

Finally, we included studies from unpublished sources, such as dissertations and theses in an attempt to reduce the chances of a publication bias. Nevertheless, most of these unpublished sources were significant. However, this does not exclude the 'file drawer effect' where non-significant findings are likely to not be published. It is also possible that the Egger Regression method is capturing other types of bias, for example the heterogeneity between the studies themselves, which is expected due to the ASD population being heterogeneous (Sterne et al., 2011).

3.5.6 Conclusions and Future Directions

Overall, it appears that individuals with ASD show lower performance measures than NT individuals on tasks involving the detection and interpretation of BM. However, age and the type of paradigm used have a great influence on the size of the difference between ASD individuals' performance and the performance of NT individuals. We show that there is a developmental delay in BM understanding, which improves with age within the ASD population and explains part of the high variability in the results established in the literature. Moreover, autistic individuals show consistently lower performance in paradigms requiring the extraction of emotion from BM in comparison to action recognition or simple BM detection. This finding is more meaningful, considering that a main characteristic of ASD is an impairment in social communication and that interaction and emotional portrayal of biological motion has great social relevance. Finally, we find that there appear to be differences between ASD and NT groups in brain activations when viewing BM and those differences can provide an insight to why the behaviour that we observe exists.

For the field of research to move forward, methodological standards need to be imposed in terms of the age ranges incorporated, and the types of paradigms used. However, interpretation standards need to be considered as well. Although it appears that there is variability in the literature as to whether and how large the effects are, the effects are actually varied due to the combination of various factors. For proper interpretation of the field, the paradigm used, and the age of the participants need to be considered as segregating factors. This is important because a child with autism might have difficulty perceiving biological motion, but by the time they reach adulthood, that effect might have subsided. Similarly, individuals with autism might find it much more difficult to extract emotion information from human movement, but they are much better at describing non-affective actions. Finally, as a field, autism research is going to find heterogeneous findings, due to the innate variability between autistic individuals. However, sound methodological principles when developing studies, will reduce that variability and allow for better consistency and easier interpretation.

Chapter 4 Are the Effects of Attention and Predictability Affected by Levels of Autistic Traits? Results from Two Dynamic Online Experiments.

4.1 Abstract

Some predictive coding theories of autism suggest that the difficulties observed in the autism phenotype are a product of imbalance of precision-weighting between bottom-up and top-down signals. One proposed mechanism for precision-weighting is attention. In two online experiments, we assessed the effects of attention and predictability using biological (BM) or coherent motion (CM) as the attentional cues. Predictability was modulated block-wise by explicitly instructing participants that 75% of targets will appear on one side of the screen, or that there is no prediction. Thus, expected targets appeared on the side congruent with the block-wise cue, and unexpected on the opposite. Performance was also correlated with scores on the short Autism Quotient (AQ) questionnaire to observe how the level of autistic characteristics affects attention and expectation. In the BM experiment (N=70), Attended targets were detected faster than unattended. Further, expected targets were detected faster than Unexpected targets, and Unexpected targets were detected slower than when expectation was set at chance. In the CM experiment (N=72), predictability effects were the same as in the BM experiment. In both experiments, there was no interaction between attention and predictability. AQ did not significantly affect performance in either experiment. Nevertheless, there were consistent but weak effects of AQ, suggesting special treatment of unpredictable but Attended targets. These results speak in favour of theories of autism that suggest that prediction errors are weighted higher, but this is dependent on the effects of attention.

4.2 Introduction

The predictive coding framework argues that the brain attempts to maintain homeostasis and minimise entropy by making predictions about the incoming information from the environment (Friston, 2009; Sajid et al., 2019). When a prediction error is encountered, the system can either update the predictive

model, or act out on the environment to change it. Learning in the predictive processing framework occurs only if enough weight is given to the incoming information that contradicts the original prediction. Thus, variabilities between individuals and in turn in some neurodevelopmental conditions can be seen as differences in this general neurocognitive mechanism. Specifically, some predictive coding theories of autism suggest that the characteristics observed in the phenotype of autism spectrum disorder (ASD) are a product of imbalance of precision setting of predictions about the world and the prediction error (Lawson et al., 2014; Van de Cruys et al., 2014, 2017).

Van de Cruys et al. (2014) propose that there is inflexible high precision setting of prediction errors in autism leading to the development of very precise predictions, which in turn create more prediction errors. They termed their theory the High Inflexible Precision to Prediction Errors in Autism (HIPPEA) theory. Later research indicates that autistic individuals overestimate the volatility of the environment, which could lead to the differences in precision setting (Lawson et al., 2017). Under the assumption that individuals with autism would learn based on prediction errors, Van de Cruys et al. (2017) argue that it is possible that a higher-level prediction is established that creates the expectation of a very variable environment. However, as expanded on by Palmer et al. (2017), differences in the ability to form a prediction about the uncertainty of an environment appear to be one of the more consistent findings, where in volatile environments, individuals with autism tend to put more weight on prediction errors and the sensory information. Thus, Palmer et al. (2017) argue that it is unreasonable to argue that the high precision setting of prediction errors is inflexible and consistently set high as proposed by Van de Cruys et al. (2014, 2017). Instead, Palmer et al. (2017) argue that precision setting is context dependent in autism. This argument is also supported by the majority of literature that has investigated predictive coding in autism (Goris et al., 2018; Lawson et al., 2017; Robic et al., 2015). Nevertheless, there is research that suggests that autistic individuals also show heightened learning rates in stable environments (Crawley et al., 2019). Hence, the disagreement between Palmer et al.'s (2017) and Van de Cruys et al.'s (2014, 2017) views on where in the processing hierarchy the differences between autistic and non-autistic individuals lies is still to be resolved.

It has been suggested that attention is what modulates the precision setting of the prediction error in estimating whether the prediction error should be discounted as due to noise or not (Feldman & Friston, 2010; Parr & Friston, 2019). This idea has been supported by neuroimaging research like the one by Kok, Rahnev et al. (2012). They cued the attention of participants to one side of the screen, while at the same time creating an expectation about the probability of that stimulus occurring on either side. fMRI results showed that when an Unexpected target was presented at an attended location there was lower activation in the primary visual cortex in comparison to when an expected target was presented at the attended location. Additionally, they showed that attention amplifies the activation to expected stimuli. On the other hand, unattended expected stimuli show a decreased activation in comparison to unattended Unexpected stimuli, which is consistent with the entropy minimisation hypothesis of the predictive processing framework (Friston, 2009). In this way, attention reverses the effects of habituation and allocates more precision to both the prediction error and the predicted content, as both are more relevant and should be acted upon.

Since precision setting can be associated with attention, it is necessary to explore the attentional properties in autism, while orthogonally modulating expectation and attention in the same paradigm. However, as in almost every other field of research in ASD, attention research shows varying results, which is to some extent due to the variety of methodologies used (Ames & Fletcher-Watson, 2010; T.-C. Chen et al., 2020; Keehn et al., 2013; Orekhova & Stroganova, 2014), and potentially due to the variability within ASD itself (Boxhoorn et al., 2020; Gargaro et al., 2018). It has been suggested that in passive tasks, autistic individuals tend to show a reduced mismatch negativity (MMN) component in comparison to neurotypical individuals, indicating lower surprise (T.-C. Chen et al., 2020; Dunn et al., 2008; Keehn et al., 2013; Schwartz et al., 2018). On the other hand, the reduced activation in neural components associated with automatic attentional shifts to novelty (i.e., MMN and P3a components) is not seen when the task is active and the stimuli are attended (Orekhova & Stroganova, 2014). Therefore, in accordance with St. John-Saaltink et al.'s (2015) findings about attentional load and MMN components, it is possible that attentional resources are reduced or at least difficult to allocate in

autism. Therefore, not spontaneously directing attention to stimuli might be the underlying mechanism in individuals with autism, rather than an inability to reorient to novel stimuli (Orekhova & Stroganova, 2014).

Thus, to be able to perform a task and explicitly modulate attention and expectation, the stimuli, to which a response is required, must be part of an active task to avoid any confounds. A paradigm that can vary expectation and attention, is the one used by Kok, Rahnev et al. (2012). Their paradigm is a variation of the Posner cuing paradigm, with the incorporation of an expectation component. In their task, attention was directed explicitly by telling participants to respond only to the left or the right-side appearing stimuli. Whereas this is informative in an fMRI design, in a fully behavioural paradigm, this would not be informative. If we explicitly direct attention to only some of the stimuli, we would not be able to get behavioural response to unattended stimuli. Instead, as the HIPPEA theory argues that the prediction error is inflexibly weighted higher than the prediction – i.e., the precision is always weighted higher for the prediction error - we would expect that covert as opposed to overt attention should be the studied mechanism. This would be caused by stimuli in the periphery appearing more salient, thus capturing covert attention even when overt attention is directed at a different location. If attention is already covertly deployed to one or the other side of a screen by the central stimulus, then the effect of expectation should provide an additional level of facilitation. However, with overt attention, the combination of expectation and attention will not be visible if attention is overtly directed at one side of the screen.

To covertly direct attention, here we modified the paradigm developed by Shi et al. (2010) which utilises a biological motion stimulus– i.e., a point-light display (PLD). The PLD acted as a covert attentional cue which reflexively directed attention to the walking direction. Furthermore, following research by Gervais et al. (2010) which showed that implied throwing action through a static image best directed attention at short (100ms), medium (300ms) and long (500ms) stimulus onset asynchrony (SOA), we used an equally directional stimulus (kicking). To further introduce an expectation component, we utilised the modification used by Kok, Rahnev et al. (2012). In this sense, participants were told where on the screen to expect a target (a Gabor patch) to appear with three possibilities: 75% chance of the target appearing on the left, 75% chance of

the target appearing on the right, and no prediction, where the target is equally likely to appear on the left and on the right of the screen. By introducing the expectation explicitly, we would avoid the issue of not all participants picking up on the statistical regularity. This arrangement of the paradigm provides us with orthogonal modulation of both attention and expectation.

To avoid potential issues with attention not being properly driven by the biological motion in individuals with high autistic traits (M. A. Atkinson et al., 2018; Orekhova & Stroganova, 2014; Todorova et al., 2019; Y. Wang et al., 2018), we ran a second experiment, where instead of biological motion we used a circle of coherently moving dots as the central attention stimulus. In this way we can check whether any effects are caused by the modulation of predictability and attention, or by the type of cue.

The general expectations in this study follow findings by Kok, Rahnev et al.'s (2012) but translated for behavioural rather than brain imaging performance where unpredictable and unattended events should show slower response time than predictable and unattended events. Predictable attended stimuli should show faster reaction time than the unpredictable attended ones. This would come under the expectation that a prediction will be favoured more, than the prediction error, as most of the time – i.e., 75% of the time, the prediction will be accurate. We expect to observe two main effects: a main effect of attention, where attended stimuli are detected faster, and a main effect of expectation where expected stimuli are observed faster.

With the addition of autistic traits, we would expect a different distribution of the responses. Specifically, we would expect the difference between the unpredictable and predictable attended stimuli to be smaller if not in the opposite direction with higher autistic traits. This will be true, if we accept that attention focuses precision on prediction errors to a larger extent in autism. As Orekhova and Stroganova (2014) propose that because of limited attention allocation in autism, attentional resources would be difficult to direct to uncued stimuli, we do not make any specific prediction about the unattended condition and the variability within it. However, if attention allocation is not an issue, then according to HIPPEA we would expect unattended unpredictable targets to also be weighted higher, with faster detection in individuals with high autistic

traits than in those with low autistic traits. Unattended and predictable targets, however, should not show a difference from individuals with lower autistic traits, as numerous studies have shown that learning and the creation and utilisation of priors is not affected in autism (Palmer et al., 2017). Finally, in the 50/50 condition we will also expect to observe the individuals with higher autistic traits to show faster reaction times, as in that situation, the uncertainty of the environment would encourage individuals to rely more on the incoming stimulus (Palmer et al., 2017; Van de Cruys et al., 2014). However, it is important to investigate the effect of attention in this state.

4.3 Methods

4.3.1 Participants

A total of 200 participants took part in the online studies. To collect a minimum of 40 students for each experiment, data was collected until there were 40 participants whose data was complete or complied with our data quality criteria (see Analysis section). One hundred and nine participants were recruited for the biological motion experiment and 91 for the coherent motion one. Forty of these participants from each experiment were recruited from Prolific.co and were paid £3.25 for their participation. The rest were recruited from the undergraduate Psychology population and were given participation credits for their time. Ethical approval for the conduct of these experiments was obtained from the University of Glasgow College of Science and Engineering Ethics Committee. All participants were provided with an information and consent screen, to which they had to actively progress through. All participants were provided with debriefing information at the end of the experiment. The undergraduate students received an extended debriefing sheet to facilitate their learning experience. The age range of Prolific participants was set to match the age range of the undergraduates for consistency. Participant demographics as well as information for the browsers and their operating systems for the samples used in the analysis can be seen in Table 9.

Table 9. Participant characteristics for the biological motion (BM) and coherent motion (CM) experiments

Experiment	Participant characteristics					Operating System (N)			Browser (N)			
	Sample	N	Age (SD)	Sex M/F/ Other	AQ (SD)	Mac OS	Win	Lin	Mozilla Firefox	Google Chrome	Safari	Other
BM	Prolific	30	21.9 (2.32)	20/10	3.96 (1.67)	1	29	0	5	23	0	2
	Undergrads	40	18.5 (1.48)	3/37	3.53 (1.84)	16	24	0	6	28	2	4
	Total	70	19.96 (2.53)	23/47	3.71 (1.77)	17	53	0	11	51	2	6
CM	Prolific	32	20.44 (1.81)	16/16	3.53 (1.84)	1	30	1	4	25	0	3
	Undergrads	40	18.75 (1.24)	4/35/ 1	3.02 (1.47)	21	19	0	2	31	2	5
	Total	72	19.5 (1.73)	20/51 /1	3.25 (1.65)	22	49	1	6	56	2	8

4.3.2 Stimuli

Stimuli were created using Psychtoolbox v3.0.16 (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) through MATLAB R2019b (MathWorks, 2019)

4.3.2.1 Biological motion

Biological motion stimuli were created through the Biomotion Toolbox (van Boxtel & Lu, 2013). The videos consisted of a kicking PLD represented with white dots [RGB: 256, 256, 256], on a grey background [80x80x80]. The 3D motion coordinates were taken from an already existing set of PLDs. The original video consisted of 360 frames recorded at 30fps. Only the first 120 frames were chosen and from those the presentation was down sampled to every 2nd frame creating a total of 60 frames at 30 fps in mp4 format. The videos were then cropped to 1s representing only the kicking motion (without the preparation for the kick) and sped up to twice the speed to create a presentation of 500ms at the most common screen refresh rate of laptops of 60Hz using Windows's Video editor application. This resulted in 15 frames at 30 fps. Videos were created at 512x512pix, however, Windows' Video editor application resized them to 540x960pix. Two videos like this were created – one kicking to the left and one to the right.

4.3.2.2 Coherent motion

The coherent motion videos consisted of white dots [RGB: 256, 256, 256] in a circular aperture [300x300pix], on a grey background [80x80x80]. The dots were presented in a 280x280pix circular aperture in the centre of the screen. Dot size was 7pix and they were randomly assigned a position. A video with leftward moving dots was created with 60 frames at 30fps in mp4 format. The videos were cropped to 1s and sped up to twice the speed to create a presentation of 500ms at the most common screen refresh rate of laptops and home screens of 60Hz using Windows' Video editor application. This resulted in 15 frames at 30 fps. Videos were created at 512x512pix, however, Windows' Video editor application resized them to 540x960pix. The dots were in the centre. Two videos like this were created – one with the dots moving to the left and one with the dots moving to the right.

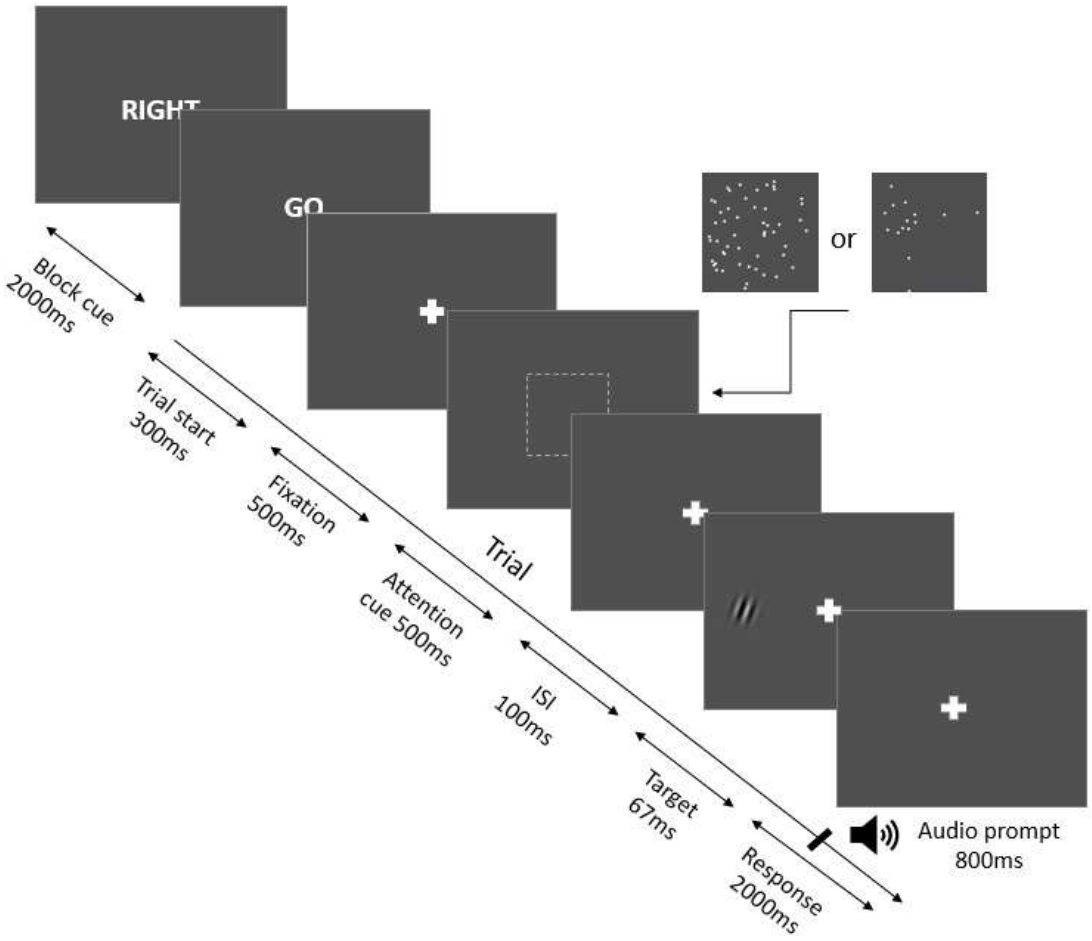
4.3.2.3 Gabor patch

The Gabor patches were set at 90% contrast with 5 cycles and 0 phase. The patch was created on the same grey background as the rest of the stimuli [80x80x80]. The images were cropped at 251x251pix, with the Gabor patch in the middle at 250x250pix. There were two patches created – one tilted to the left by 20deg (counter clockwise CCW) and one to the right by 20deg (clockwise, CW).

4.3.3 Paradigm

The experiment was created using PsychoPy3 v2020.1.3 (Peirce et al., 2019). A graphical representation of the paradigm can be seen in **Error! Reference source not found.**. The experimental set up was created for participants sitting 60cm from the screen. Background colour was set to match the background of the biological/coherent motion and the grating stimuli, which was done using a colour matcher and set to [-0.380, -0.380, -0.380] and RGB colour space. All instructions were white [1, 1, 1], font Arial, and their size varied based on the amount of text to allow it to fit on the screen.

Figure 11. Graphical representation of the paradigm. The Biological motion experiment used a kicking PLD as the attentional cue (right), whereas the coherent motion experiment used a circle of coherently moving dots as the attentional cue (left).



Following the presentation style of Shi et al. (2010), and the trial structure of Kok, Rahnev et al. (2012), we utilised the following block structure. Each block started with a 2000ms screen indicating what the expectation should be (left, right or none). This was followed by a fixation cross [size – w:0.5, h:0.5][3] lasting 500ms. Then the biological motion stimulus [0.8, 0.5] or the coherent motion stimulus [0.8, 0.5] were presented for 500ms, followed by an interstimulus interval of 100ms with a fixation cross. Next, a grating [0.1, 0.1] was presented on the left or right (x=0, y=± 0.4) side of the screen for ~67ms. Participants had to respond as fast as possible to the orientation of the grating stimulus (CW or CCW) by pressing either the B or the H keys. This choice of keys was made as we observed a Simon effect while designing the experiment, where fast erroneous

[3] The native measurement 'height' was used to distribute the stimuli across the window on which the experiment is open – set to automatically open in full screen. This allowed PsychoPy to use the window size and resolution of the participants' own screens. The sizes of the stimuli reported in the paragraph correspond to the measurements given to PsychoPy in 'height' unit.

responses were made by pressing the key located on the left side of the participant when the target appeared to the left, rather than as a response to target's orientation. Thus, choosing the B and H keys allowed some vertical separation. The next trial started after the participants' response or 2000ms after the disappearance of the grating stimulus. After 800ms post target disappearance, if the participants did not respond, they heard a short – 200ms, tone (220Hz – A-tone) prompting them to respond. A new trial was indicated by a short screen (300ms) showing the word GO. In previous experiments in our lab, participants were showing difficulty in tasks that have continuous trial structure, thus the visual break was introduced.

To maximise participant engagement online we had a total of 30 blocks (no longer than 21min total). Each block had 12 trials; a total of 360 trials – 180 trials per attentional cue level (left/right). At 75% predictability (9/12 congruent with block cue), this produced 180 [90/90 attended/unattended] expected trials and 60 [20/40 attended/unattended] unexpected. Additionally, there were 120 [60/60 attended/unattended] trials at the 50/50 condition. The central cue was congruent with half of all presentations in each expectation condition, as the central cue had a predictability of 50% across blocks.

4.3.4 Procedure

Before the task, participants were asked for their age, sex, eyesight, English language fluency, as well as their browser and operating system, as these can contribute to variability in response times (Anwyl-Irvine et al., 2020; Bridges et al., 2020).

At the Information sheet screen, participants were asked to increase their brightness to 100% and to stay approximately at 60cm from their screen. The experimental task started by firstly familiarising participants with all the components of the task – the Gabor patches at the corresponding response buttons; the block cues – left/right/none; the biological/coherent motion stimuli. Participants were also asked to adjust their volume, so they can comfortably hear the beep. After familiarising themselves with the task, participants performed 8 trials with feedback to get them used to the task, without a predictive word at the beginning. Participants were informed that

they need to respond as fast and as accurately as possible to the orientation of the Gabor patch and that if they do not answer within 800ms they will hear a beep to prompt them to respond. They were also told that the central cue was not predictive of the location of the target. Participants were allowed to take a break between the blocks, and they were provided with a counter of how many blocks they had left of the task. At the end of the task, participants were asked to state what was their task, to make sure that they performed according to task instructions. Additionally, they were asked how well they thought they performed on the task. Finally, as the task was performed online, sometimes refresh rate issues interfered with video presentations. For this reason, we asked participants how many kicks they saw per trial for the biological motion video.

Finally, participants answered 10 questions comprising the short version of the autism quotient questionnaire (AQ10, Allison et al., 2012).

4.3.5 Power

Since we have only effect size information about the effect of the attention cue from Shi et al. (2010) d=1.14, and the effect of expectation from Kok, Rahnev et al. (2012) – d=0.79, we used the PANGEA applet with a 2(attention)x3(expectation) design to estimate sample size for the main effects and interactions. As there is variability in the number of trials each condition combination gets, we used the smallest number as a representative for all of them – 20 trials per participant for the Unexpected Attended condition per congruency. Also following suggestions from Brysbaert (2019), the expected effect size was lowered to 0.4 as more realistic for both the main effects and the interaction. With 40 participants, d= 0.4 for main effects and interaction, and 20 replicates, we will get power above 81%. As PANGEA does not provide the option for adding a continuous variable to the calculation, we decided to add the AQ sores as a categorical variable, with participants nested within the variable. We gave this variable 2 levels, with 20 participants per level, assuming that the AQ can be split in low and high. Again, with 20 replicates and d=0.4, all main effects and interactions indicated power of above 80%, except the main effect of AQ, which reaches 69% power and the Predictability*AQ interaction, which reaches 57%. However, as we are not going to be splitting participants in half, we did not consider the lower power as indicative as splits of continuous

variables lead generally to lower power (McClelland et al., 2015) and we were going to use the AQ as a continuous variable. Nevertheless, we decided to recruit 40 students per study with additional samples recruited from Prolific, thus increasing power and the possibility to detect an effect. Whereas Brysbaert (2019) shows that for an interaction effect and one post-hoc in a 2x2 design you would need ~100 participants, this is related to one observation per participant. Although we have an imbalance in the design, the larger number of replicates allow us to recruit a smaller sample.

4.3.6 Analysis

4.3.6.1 Data quality

As data were collected online, participants' performance is sensitive to an individual participant's set up. Thus, the log files were used for accurate timing of responses, to identify participants who did not follow the instructions or whose set up introduced too long delays in stimuli presentation. Two types of criteria were imposed to maintain data quality – response wise and participant wise. Response-wise criteria included removal of responses below 300ms or above 1500ms after target onset and keeping only the first responses that occurred 50ms after target onset. Participant-wise criteria for removal included accuracy below 50% (*N*=10), more than 30% too slow or too fast responses (*N*=8), wrong answer on the comprehension question and an accuracy below 65% (*N*=7), pressed a response button more than twice per trial (to allow for accidental button presses) on average (*N*=0), participants indicated that the PLD was presented mostly twice (i.e., two kicks) each trial (for BM study only) (*N*=5). Additionally, due to observed delays in target displays, a window of 50-100ms was chosen based on the presentation times in Kok, Rahnev et al. (2012) and Shi et al. (2010) – if more than 50% of targets were presented outside this window, the participant was removed (*N*=28).

If the participants answered incorrectly to an attention check placed in the AQ10, their AQ score was not calculated, and was replaced with the mean score of the participants from the platform they came. The imputation was done for eight participants in the CM experiment and eight for the BM experiment.

All data, experiment design and analysis scripts will be shared on ReShare.

4.3.6.2 Statistical analysis

We analysed the raw reaction time responses using generalised mixed effects models, with participant and target grating (left, right) as random effects.

All analyses were done using R(v4.0.4) (R Core Team, 2021) and RStudio (v1.3.1093) (RStudio Team, 2020). Models were build using the lme4 (v1.1-27) package (Bates et al., 2015), post hoc analysis were done using the emmeans (v1.6.1) package (Lenth, 2021). All data wrangling was done using tidyverse (v1.3.1) and attached packages (Wickham et al., 2019).

Since we are modelling reaction time (RT) data, which is bound to always have a minimum value of zero and is therefore most commonly right-side skewed, we used a Gamma identity model family to estimate the coefficients from a generalised mixed-effects model. Attention (Attended/Unattended), Predictability (Expected, Unexpected, None [50/50 case]), Platform (Prolific/student) and AQ were the fixed factors. The former three were deviation coded, and the AQ was z-score standardised. The targets – left- and right-tilted gratings, participant IDs and their interaction were used as random factors. Radom intercepts were estimated for all random factors. Slopes were estimated for the main effects and interaction for the participant X target tilt random factor, and slopes for the main effects were estimated for the participant random factor. This structure was chosen based on principal components analysis on the random effect structure, performed to deal with singularity as recommended by Bates et al. (2018). Models were compared with the maximal model, and no statistical differences were observed (p-values > 0.8).

Main effects and interactions were decomposed. Tukey adjustment was used for multiple comparisons of the p-values and the confidence intervals (CI) for the Attention factor contrasts and Sidak adjustments for the Expectation factor contrasts and pairwise comparisons as recommended by emmeans package.

The results from the BM experiment were also replicated in a second Prolific sample. The results can be found in the Appendix D.

4.4 Results

4.4.1 Biological motion

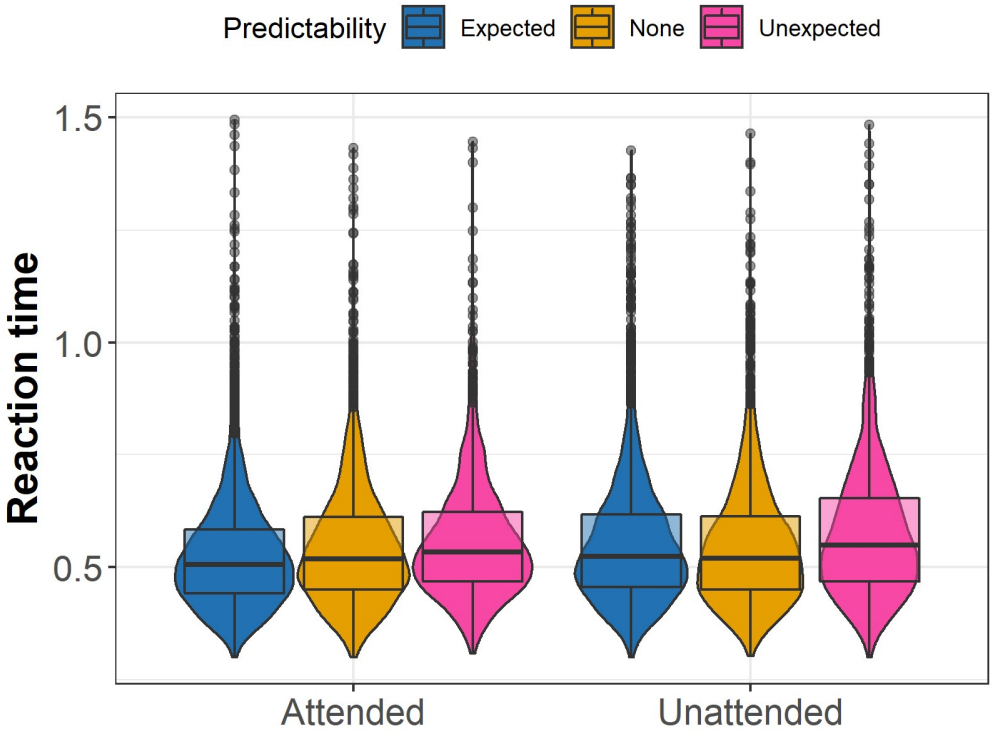
Reviewing the summary statistics in Table 10, variations between different conditions are small for accuracy, indicating that participants were able to perform the task. Further, reaction time (RT) was faster for Expected targets in comparison to Unexpected targets. In the None condition, reaction time to Attended and Unattended targets are equivalent, although accuracy was higher in Attended condition. This difference can also be observed in Figure 12.

Table 10 Participant accuracy and median reaction time in seconds.

Attention congruency	Prediction congruency	Mean accuracy (SD)	Median RT (MAD)*
Attended	Expected	0.928 (0.259)	0.502 (0.106)
	None	0.917 (0.277)	0.517 (0.122)
	Unexpected	0.893 (0.309)	0.532 (0.120)
Unattended	Expected	0.920 (0.271)	0.519 (0.120)
	None	0.911 (0.286)	0.517 (0.122)
	Unexpected	0.903 (0.296)	0.543 (0.135)

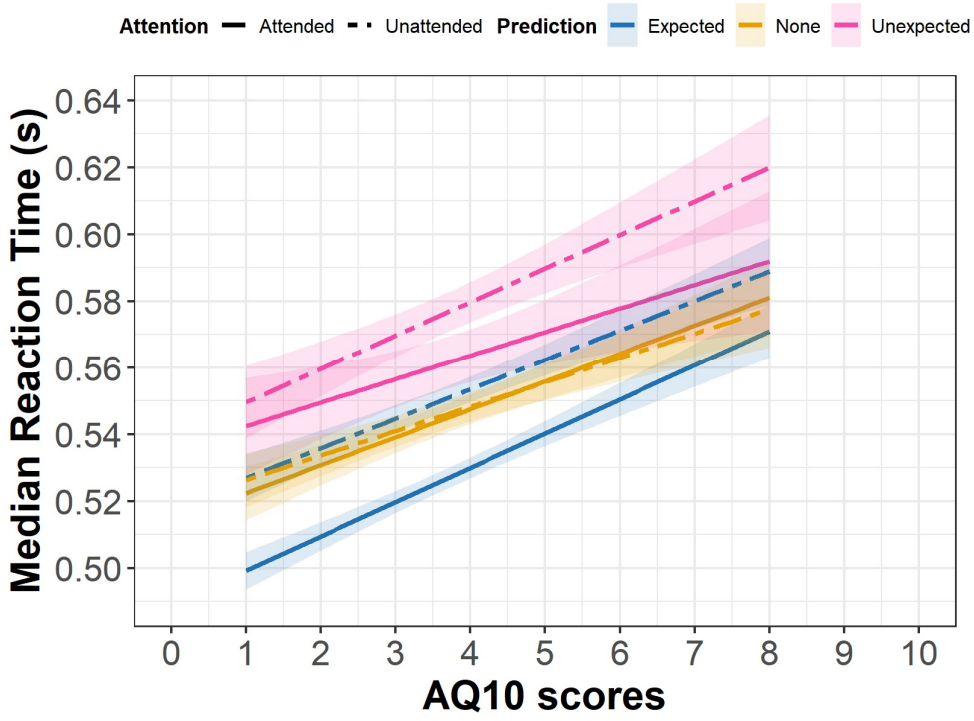
***Note*: * MAD = Median absolute deviation**

Figure 12. Reaction time in seconds for accurately detected targets by attention and predictability conditions.



We also observe a change in performance with change in AQ scores, mainly an increase in RT with higher AQ scores (see Figure 13). Consistent with the data we see in Table 10, Attended targets elicited shorter RTs than Unattended targets, and this difference was consistent with increase in AQ scores. However, the difference in RT in Attended and Unattended targets appears to swap with increase in AQ when no expectation is set, although that change is very small. Interestingly, with the increase of AQ, the difference between Attended and Unattended Unexpected targets grew larger (pink solid and dashed lines). Thus, Attended but Unexpected targets were closer in reaction time to Unattended Expected targets. All of these effects are also seen in the replication Prolific sample in Appendix D.

Figure 13. Changes in RT with AQ scores for Attention (line types) and Prediction levels (colours).



These increase in RT with increase in AQ, however do not come out as statistically significant. Overall, there was a significant main effect of Attention ($X^2(1) = 17.656$, $p < 0.001$) and Predictability ($X^2(2) = 34.239$, $p < 0.001$), however, there was no significant interaction between the two variables ($X^2(2) = 5.914$, $p = 0.052$). Despite the observed differences in Figure 13, there was no significant main effect of AQ ($X^2(1) = 3.244$, $p = 0.072$), and no significant 2-way interaction with Predictability ($X^2(2) = 0.257$, $p = 0.879$) or with Attention

($X^2(1) = 0.049$, $p = 0.825$). There also was no significant 3-way interaction ($X^2(2) = 1.253$, $p = 0.535$). This suggest that the effect of AQ scores were not strong and consistent enough to influence participants' performance. There was an interaction between Attention and Platform ($X^2(1) = 4.086$, $p = 0.043$, see Table 11 below). No other effects or interactions were significant with Platform – all *p*-value*s* > 0.127, indicating that there were no other differences between the different platforms.

Overall, Attended targets were detected faster than Unattended ones – $\beta = 0.039$ [SE=0.092, 0.021 – 0.057], z-ratio = 4.202, $p < 0.001$. Additionally, as seen in Table 11, students showed a bigger difference in response time between Attended and Unattended targets, in comparison to the participants recruited through Prolific.

Table 11. Simple effects of attention at each platform level

Contrast	Platform	estimate	SE	LCI	UCI	*Z*-ratio	*p*-value
Attended - Unattended	Prolific	0.021	0.014	-0.011	0.052	1.476	0.260
Attended - Unattended	Student	0.057	0.012	0.031	0.084	4.804	<0.001*

***Note*: * indicate statistically significant contrast.**

Next, the decomposition of the Predictability effect showed that Expected targets were detected faster than None targets, however this difference was not significant - $\beta = 0.016$[SE=0.010, -0.007 – 0.039], z-ratio = 1.664, $p = 0.261$. None targets, however, were detected faster than Unexpected targets – $\beta = 0.068$[SE=0.014, 0.035 – 0.101], z-ratio = 4.930, $p < 0.001$, and Expected targets were detected faster than Unexpected targets - $\beta = 0.084$[SE=0.015, 0.050 – 0.119], z-ratio = 5.801, $p < 0.001$. Thus, as expected, predictability assists in participants ability to respond faster to expected targets, however, that ability is not significantly modulated by attention or AQ scores.

4.4.2 Coherent motion

Moving to the CM experiment, in Table 12 we see that variations between different conditions are small for accuracy indicating that individuals were able to do this task as well as the BM experiment. Additionally, RT is slightly faster for Expected targets in comparison to Unexpected targets. Further, looking at

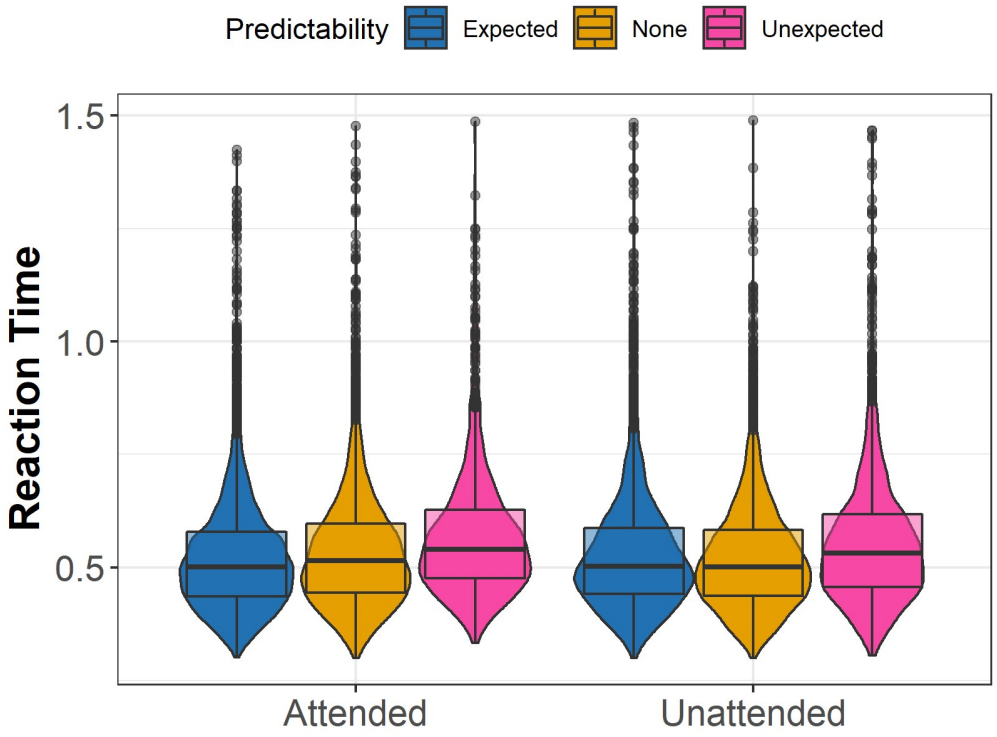
the Unexpected and the None conditions, Attended targets elicited longer reaction times in comparison to the Unattended target which is an indication of inhibition of return (IOR, Klein, 2000), which was not observed in the BM experiment. This can also be observed in Figure 14.

Table 12 Participant accuracy and median reaction time in seconds

Attention congruency	Prediction congruency	Mean accuracy (SD)	Median RT (MAD)*
Attended	Expected	0.924 (0.265)	0.500 (0.102)
	None	0.909 (0.288)	0.508 (0.112)
	Unexpected	0.889 (0.314)	0.533 (0.114)
Unattended	Expected	0.918 (0.275)	0.501 (0.107)
	None	0.908 (0.290)	0.500 (0.105)
	Unexpected	0.917 (0.276)	0.528 (0.116)

***Note*: * MAD = Median absolute deviation**

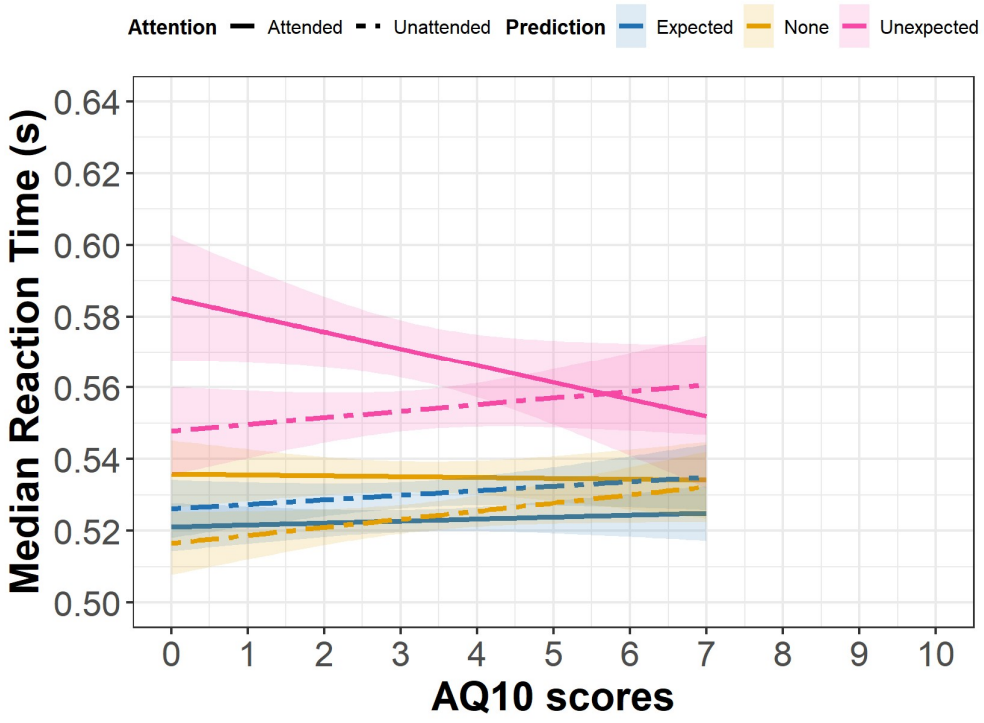
Figure 14. Reaction time in seconds for accurately detected targets by attention and predictability conditions.



Further, unlike in the BM experiment, AQ scores did not appear to show any effects on performance (see Figure 15). Attended targets led to longer RTs, however, this was the reverse for Expected targets. Thus, it appears that expectation reverses the descriptively observed IOR, by eliciting faster RTs to Attended in comparison to Unattended targets. Although this trend stays consistent with AQ, similar to the BM experiment, higher AQ leads to shorter RTs

for Unexpected Attended targets, and the difference reduces between Attended and Unattended targets when no expectation is set.

Figure 15. Changes in RT with AQ scores for Attention (line types) and Prediction levels (colours).



Unlike in the BM experiment and despite the indication of an IOR in the RT data, the main effect of Attention was not significant ($X^2(1) = 0.898$, $p < 0.343$). This is potentially caused by the observed effects of AQ and expectation seen in Figure 15.

Next, similar to the BM experiment there was a significant effect of Predictability ($X^2(2) = 56.754$, $p < 0.001$). The interaction between the two variables was also not significant ($X^2(2) = 0.830$, $p = 0.660$). Additionally, there was no significant main effect of AQ ($X^2(1) = 0.059$, $p = 0.808$), and no significant 2-way interaction with Predictability ($X^2(2) = 0.695$, $p = 0.706$) or with Attention ($X^2(1) = 2.694$, $p = 0.101$). There also was no significant 3-way interaction ($X^2(2) = 1.667$, $p = 0.435$). This suggest that AQ scores did not significantly influence participants' performance. No other effects or interactions were significant with Platform – all p-values > 0.205, indicating that there were no differences between the different platforms, unlike in the BM experiment.

Similar to the results in the BM experiment, Expected targets were detected faster than None targets, however this difference was not significant – β = 0.012[SE=0.010, -0.013 – 0.037], z-ratio = 1.122, p = 0.598. None targets, however, were detected faster than Unexpected targets - β = 0.099[SE=0.015, 0.064 – 0.134], z-ratio = 6.685, p < 0.001, and Expected targets were detected faster than Unexpected targets - β = 0.111[SE=0.015, 0.075 – 0.147], z-ratio = 7.331, p < 0.001. Thus, as expected, and despite the potential for IOR, predictability assists in participants' ability to respond faster to Expected targets, however, that ability is not modulated by attention or AQ scores.

4.5 Discussion

In this study, we aimed to separate the effects of attention and predictability on detection speed in two modified Posner paradigms involving BM and CM. Further we aimed to see how these effects differ with autistic traits. In the BM experiment, we observe a general effect of attention and predictability, which indicate faster processing for Attended and for Expected stimuli. More specifically, predictability improved performance over situations when there was no expectation set (None condition), and hindered it, when the target was Unexpected. In the CM experiment, we did not observe a significant effect of attention. However, there was an overall effect of predictability. This mirrored the effect in the BM experiment showing that Unexpected targets delayed the participant's ability to detect them in comparison to Expected ones. Like in the BM experiment, predictability did not significantly improve performance above the None condition when comparing it to the Expected condition. Finally, and more importantly, autistic traits did not show a statistically significant interaction with any of the factors, suggesting that in this paradigm, predictability and attention are not significantly affected by the level of autistic traits. However, there were descriptive effects that suggest different treatment of Attended Unexpected targets with increase in AQ. We also see that unlike in the CM experiment, in the BM experiment individuals with higher autistic traits had longer RTs suggesting that the task was more difficult (Todorova et al., 2019).

In respect to the effects of autistic traits, from the point of view of HIPPEA, we would have expected smaller differences in RTs with increase in AQ between

Unexpected and Expected targets. We would also have expected this effect to be mostly evident in the Attended condition, with attentional resources being difficult to allocate in the Unattended condition (Orekhova & Stroganova, 2014). Unfortunately, we do not observe any significant interactions with increase in AQ. However, we see a trend in the data in both the BM and the CM experiments where Attended Unexpected targets are detected faster than the Unattended Unexpected targets, bridging closer the Expected and Unexpected Attended targets. Thus, we see some support for our prediction of smaller difference between Expected and Unexpected Attended stimuli with increase in AQ. This suggests that attention focuses precision on the prediction error to a larger extent in high AQ individuals, thus providing some descriptive support for the HIPPEA theory. However, we do not see Unattended Unexpected targets being weighted higher - we see increased RTs for Unattended Unexpected targets. This suggests that, as argued by Orekhova and Stroganova (2014), attentional resources might be more scarce in autistic individuals, and by extension the higher AQ participants in our sample, although these effects might be attenuated or qualitatively different in the general population as opposed to individuals with an ASD diagnosis. This also suggests that the 'inflexibility' part of the HIPPEA theory might not hold and although prediction errors might be given higher precision, those prediction errors would need to be attended.

Looking at attention allocation specifically, we observe an interesting pattern with the increase of AQ in the two experiments. In the BM experiment, as Unattended targets have longer RTs than Attended ones, we can safely say that Unattended targets are indeed targets where the attention has been driven to the opposite location. Thus, the descriptive decrease in RT in individuals with higher AQ for Attended Unexpected targets allows us to assume that attention is highlighting the importance of the prediction error to a larger extent for high AQ participants. However, in the CM experiment where we descriptively observe IOR, we still observe the same decrease in RTs with increase in AQ for Attended Unexpected targets. Moreover, this decrease crosses over and improves on the RT for Unattended Unexpected targets, which in an IOR circumstance should be detected faster. This suggests that attention might be differently allocated in individuals with higher AQ traits. To further understand this pattern and taking a more speculative look at the descriptive results, we will next examine this

effect by reviewing the differences in RT for Attended and Unattended targets with increase in AQ in the None condition, where no expectation is set.

In both the BM and CM experiments, we see a decrease in the distinction between Attended and Unattended targets when no expectation is set in the None condition. Although it is possible that the BM cue was not producing large attentional effects, there appears to be a greater distinction between Attended and Unattended targets in the CM experiment for all participants. Thus, since we see the decrease in the distinction between the two attention levels with increase in AQ in the CM condition as well, it is possible that individuals with higher AQ traits are less influenced by the attentional cues in the absence of a prediction model in the None condition. On the other hand, in the blocks where expectation was set, attention starts to influence the behaviour of individuals with higher autistic traits. Interpreting this through the predictive coding models of autism, these descriptive results indicate that with the increase in AQ, when no expectation is set and in a volatile environment, these individuals tend to rely more on the incoming sensory information (Lawson et al., 2014; Van de Cruys et al., 2014, 2017). However, when a predictive model is established, that model will be used to interpret the environment. Further, in cases of IOR it is argued that if high precision is allocated to the volatility of the environment – i.e., we expect that the state of the environment could change we are more likely to experience shorter IOR (Parr & Friston, 2017). Thus, it is possible that our participants with high AQ were overestimating the volatility of the environment and thus they were more likely to attend to previously attended locations. This latter interpretation is more consistent with Palmer et al.'s (2017) and Lawson et al.'s (2017) argument about precision setting being context dependent in autism, and more closely related to the volatility estimation, rather than being set 'inflexibly' high as suggested by HIPPEA.

As mentioned, these interpretations however are speculative and the effects need to be replicated. Hence, the only component that the present results can support from the predictive coding perspectives and previous studies is that individuals with higher AQ scores are able to use statistical regularities to guide their behaviour as well as their lower AQ counterparts (eg., Allenmark et al., 2020; Chambon et al., 2017; Van de Cruys et al., 2018). Thus, despite the predictive coding literature arguing that ASD is a product of imbalance of

precision setting (Lawson et al., 2014; Van de Cruys et al., 2014, 2017) and that differences will be most pronounced in unpredictable environments (Palmer et al., 2017; Van de Cruys et al., 2014), we were not able to provide concrete support for these arguments.

It is also important to note, that in these experiments, our participants were individuals from the general population and the conclusions we are making about the utility of HIPPEA and the other predictive coding theories of autism are based on the participants' AQ scores. Thus, although autistic traits vary within the general population (Ruzich et al., 2015), any effects would not be as strong as the ones observed in autistic individuals. The descriptive results that we see with the change in AQ, could be more prominent in an autistic sample. This becomes evident with the combination of the replication and main samples (see Appendix D), where the larger sample size, pushed the effect of AQ over the chosen significance threshold. However, it is also possible that we observe a qualitatively different result in autistic individuals. For example, we could observe that both Attended and Unattended Unexpected events are detected faster by autistic individuals and in this way differentiating neurotypical individuals with higher autistic traits from those with a diagnosis. In this way the inflexible part of HIPPEA might still be an appropriate suggestion. Hence, further replication of these results within an autistic sample would be necessary to provide more concrete support for the present findings and evidence for or against the predictive coding theories of autism.

From a replication point of view, similar to the results by Kok, Rahnev et al. (2012) in the BM paradigm we observed a highlighting effect of attention, where attended stimuli were detected faster than unattended for both Expected and Unexpected stimuli. These results are also concordant with results showing that attention amplifies the neural signal (Mehrpor et al., 2020). However, this was not the case for the CM experiment. From a behavioural perspective, we observed what descriptively looked like IOR. It has been previously shown that BM stimuli are difficult to ignore and they are incidentally processed (Thornton & Vuong, 2004), thus potentially capturing attention for longer, rather than leading to disengagement with the cued side. However, we believe that the descriptively observed IOR at the CM experiment was probably caused by the different motion properties of the PLD and the CM stimuli. Whereas in the PLD

the action is concentrated at the end of the 500ms (the human first leans and then kicks), in the CM stimulus the directional motion is present throughout the 500ms. Additionally, it has been suggested that it takes longer to integrate the information in a BM stimulus than in simple motion stimuli (Neri et al., 1998). This could have also contributed to the expected attentional guiding processes in the BM experiment in comparison to the CM. A model combining the BM and CM experiments (see Appendix D) also shows that indeed the BM was better at capturing attention, leading to larger differences between Attended and Unattended trials, as opposed to the CM stimulus. Importantly, there was no interaction with AQ scores. Due to these differences, we do not believe the effects were caused by the type of cue (social/non-social) but by the motion properties of the cue.

Despite the descriptively observed IOR, we also saw that expectation overwrote the effects of attention leading to indistinguishable differences in the detection of Attended and Unattended Expected targets. This, although non-significant, was potentially the reason for the absence of a main effect of attention in the CM experiment. This interpretation suggests that predictability can override the effects of attention. Kok, Rahnev et al. (2012) argue that attention facilitates expectation, thus, since descriptively we observed IOR, RTs for Attended Expected targets should have been slower. Instead, by observing the opposite (faster RT for Attended Expected targets in comparison to Unattended), we argue that attention takes a backseat and expectation drives performance. As suggested by Parr and Friston (2017), expectation in this setting could be increasing the expectation for the sensory information to change and in this way reversing/minimising the IOR effect for participants. This reversal of attention at the expected condition suggests that although attention acts as a highlighter of all events regardless of expectation status, expectation acts as a driving force, even when attention is actively biasing behaviour in the opposing direction. This difference is important, as it reaffirms that predictability is salient (Southwell et al., 2017; Yon et al., 2020), despite attentional demands. However, it needs to be taken into account that these differences are again not significant and that we observed slightly different effects in the student and the Prolific samples even in the BM experiment, with Prolific samples descriptively showing an IOR when no expectation was set. Thus, it is important for these

results to be replicated in studies that investigate attention and prediction not only when attention is expected to assist but also when attention is driven away from the Expected target.

The present results and discussion need to be considered with a degree of caution due to several methodological caveats. Firstly, the data is collected online, which would increase the variability and decrease the sensitivity of the results. Specifically, while participants were told to increase the brightness of their screen, and to be 60cm away from the screen, we cannot be sure that all participants followed these instructions. However, not following these instructions would have made the experiment more difficult for the participants, and that would have been reflected in lower accuracy and higher RTs, meaning they would probably be excluded from the analysis. Secondly, the use of online measures introduces large variability in the experimental presentation itself due to the different set-ups that participants would have. This is evidenced from the largest portion of participants being removed because of delayed presentation of the targets. Although, we were able to clean out and deal with some of the noise that is introduced in this way, the presentation of the targets and in turn the ability of participants to accurately detect the targets is undoubtedly affected by these differences. Thus, for future online studies, where timing is important, the recruitment of larger samples, where more stringent cleaning can be performed is necessary. Next, the definition of attention is somewhat conflated with expectation. In attention allocation paradigms, a stimulus driving the attention is creating an expectation. The reflexive attention shifts following the kick of the PLD are in essence a result of the expectation of something occurring on the direction of motion. Thus, attentional modulation not coupled with expectation would be necessary to further study the effect of attention and expectation separately. Finally, the trials on the Unexpected side are unbalanced, which is evident in the larger standard errors around the Attended Unexpected lines. Although the uncertainty around this result is larger, the consistency between samples suggests an underlying effect that needs to be investigated further.

Taken as a whole, the present findings show that individuals with high autistic traits can use explicitly stated regularities of the environment to a similar degree as individuals with low autistic traits. Moreover, we do not see a

significant indication of enhanced precision for unpredictable events regardless of attention level, suggesting no differences in the allocation of attention with the increase of autistic traits. Thus, the present findings stand in opposition with the main argument of high precision allocated to prediction errors as suggested by HIPPEA. However, some of the observed small differences might be amplified in individuals who have an official diagnosis. We conclude that the use of paradigms that modulate attention and prediction separately and introduce different attentional effects might assist in the future.

Chapter 5 Hierarchical frequency tagging of biological motion processing: a proof-of-concept

5.1 Abstract

According to the HIPPEA theory, autistic individuals allocate higher precision to prediction errors and bottom-up stimuli. One of the proposed mechanisms through which HIPPEA suggests this happens is by allocating a disproportionate amount of attention to low-level features. One recently suggested methodology that could allow the investigation of whether attention influences the combination of bottom-up and top-down information in autism is the Hierarchical Frequency Tagging (HFT) procedure. The HFT procedure allows tagging the processing of information at both the high and low levels of the processing hierarchy, along with their combination in the brain, while modulating attention and predictability. This is achieved by tagging conceptual information at one frequency, bottom-up information at another and investigating their combination through their intermodulation components observed in the EEG signal. This chapter presents a proof-of-concept EEG study, which aims to determine whether the HFT paradigm could be applied to a dynamic stimulus – a point-light display walker. We modulated attention by instructing participants to count only one direction of walking. We modulated prediction by varying the probability of one or the other walking direction. Using the data from 3 participants, the results indicate that dynamic stimuli can be easily incorporated into the HFT procedure, and that the tagging of higher-level information is dependent on the participants' conscious perception of the stimuli. Although some of the intermodulation effects were not replicated in this small sample, the overall results suggest that this type of paradigm could be used in future work to investigate the combination of bottom-up and top-down signals in a dynamic context in autism.

5.2 Introduction

One of the newly proposed theories of autism views the condition through the prism of the predictive coding framework. In this framework, the brain creates predictions about the world, which are compared to the actual input (Friston et al., 2011). As the input never exactly matches the prediction due to noise in the

environment and the brain, this comparison produces an error signal. Depending on the level of precision associated with the prediction error, e.g., how unambiguous, and relevant to current goals it is, the prediction error will lead to an update of the prediction, or it will be discounted. In this sense, prediction errors are the basis for learning, indicating when the prediction is wrong and a new rule about the environment needs to be learned (Friston et al., 2011). Following this framework, in 2014, Van de Cruys et al. proposed the theory of High Inflexible Precision of Prediction Errors in Autism (HIPPEA). According to HIPPEA, individuals diagnosed with Autism Spectrum Disorder (ASD) are deficient in meta-learning (i.e., knowing which signals indicate something that should be learned) due to higher precision setting of bottom-up information.

One of the proposed mechanisms through which HIPPEA suggests that weighting of prediction errors is affected in autism is by allocating a disproportionate amount of attention to low-level features of stimuli – i.e., the individual features of the environmental information (Van de Cruys et al., 2014). Too high weighting of prediction errors will lead to forming future predictions that are based on noise and infrequent contingencies. The HIPPEA framework argues that placing more focus on low-level information impedes the formation of a generalisable holistic percept. Thus, individuals with ASD form narrow models about the world with very specific predictions and even small variations in the input will not fit in those predictions and will be categorised as errors (Van de Cruys et al., 2014). This contrasts with the broader models created by neurotypical individuals that can correctly discard some variation created by noise in the environment without leading to the propagation of prediction errors higher up in the hierarchy. Furthermore, according to HIPPEA, abnormal sensory perception will impede social functioning since autistic individuals will create predictions that are grounded in the low-level information instead of the abstract generalisable concepts that are necessary for social interactions (Borghi et al., 2017; Dove, 2016).

This difficulty in processing social features and interactions has been seen in recent meta-analyses showing that autistic individuals consistently show lower performance in detection and interpretation of human movement from both simplified point-light displays (PLD) and full-light displays (see Chapter 3; Federici et al., 2020; Todorova et al., 2019; Van der Hallen et al., 2019).

Specifically, the more complex the biological motion tasks became – from detection to emotion interpretation, the larger the differences between NT and ASD individuals became (Todorova et al., 2019). Moreover, differences between ASD and NT participants have been linked with reduced modulation by higher-order brain areas responsible for the interpretation of biological motion such as posterior Superior Temporal Sulcus (pSTS) (Grosbras et al., 2012; Lestou et al., 2008; Thurman et al., 2016). Hence, differences in biological motion perception could be due to reduced integration at a higher level but could also be linked to already present differences in early-on processing stages.

Research investigating HIPPEA has produced mixed results. Findings point towards less influence of context (Goris et al., 2018), higher estimation of the volatility in the environment (Lawson et al., 2017), along with higher learning rates (Crawley et al., 2019), potentially influenced by a higher weighting of prediction errors. However, most findings appear to be only partly in line with HIPPEA. For example, Sevgi et al. (2020) showed that in stable environments autistic individuals do not differ from neurotypical (NT) individuals. Differences between ASD and NT individuals appeared when environments became more volatile. This suggests that prediction errors are not weighted inflexibly high. If they were, then we would expect that unexpected variations in the stable environments would impede performance, as that would impede the creation and utilisation of a flexible prediction model. Since high precision setting of prediction errors will bring about higher learning rates, differences in performance should be more evident in the stable environment as volatile environments generally would benefit from higher learning rates (Crawley et al., 2019). Hence, more work is needed to understand the ability of HIPPEA to account for autistic characteristics. Further, the existing literature struggles to disentangle where exactly the imbalance in precision setting occurs in autism - at the lower level of information processing, at the integration of the prediction error and the prediction or at the prediction itself, and instead focuses mostly on the outcome in performance.

One way to investigate whether individuals with ASD weigh prediction errors and input higher than top-down information is to tag top-down and bottom-up processing, and to investigate their interaction. Additionally, since attention is one of the proposed mechanisms for precision-weighting of prediction errors

(Feldman & Friston, 2010), it is important that when investigating predictive coding theories, attention and predictability are controlled separately to distinguish between them.

Using a Hierarchical Frequency Tagging (HFT) procedure, Gordon et al. (2017) investigated the effects of predictability and attention on the integration of the two tagging frequencies in the brain. The HFT procedure utilises higher-order information processing tagging - Semantic Wavelet-Induced Frequency-Tagging (SWIFT) (Koenig-Robert & VanRullen, 2013), and lower-level information processing tagging - Steady-State Visually Evoked Potential (SSVEP), and allows exploring their intermodulation components (IM). The SWIFT was created by alternating between house and face images and their wavelet-transformed images and the SSVEP through a flicker of the presentation screen. Attention was modulated by asking the participants to count either the houses or the faces. Predictability was modulated by controlling the ratio of faces to houses in each block. In NT participants, the SWIFT signal is suggested to reduce with increasing predictability as less weight is given to the prediction error (Gordon et al., 2017; but see Coll et al., 2020). The IM components, on the other hand, are frequencies that are produced from the non-linear combination and interaction of the two tagging frequencies. The IM signal is considered to reflect the efficiency of integration between top-down and bottom-up signals (Gordon, Hohwy, et al., 2019). With high predictability the signal-to-noise ratio (SNR) of the IM components showed better integration indicating a larger influence of the top-down model (Gordon et al., 2017). Importantly, a recent study by Coll et al. (2020) replicated some of the original findings by Gordon et al. (2017). Coll and colleagues (2020) were not able to find modulation of the SWIFT SNR with the predictability of the stimuli, however, they replicated the increase in the IMs SNR with increase in predictability. Hence, the modulation of the SWIFT by predictability may not be a stable finding.

To further explore the relationship between the lower and higher levels of hierarchical processing, Gordon, Tsuchiya, et al. (2019) used a Multi Spectra Phase Coherence (MSPC) methodology. The authors argue that by calculating the MSPC driven by the stimulus (MSPCstim) and the MSPC that is driven by the resulting EEG response (MSPCres), we can observe the integration of bottom-up and top-down signals at different levels of the hierarchy. Gordon, Tsuchiya, et

al. (2019), and later elaborated in Gordon, Hohwy et al. (2019), implicated MSPCstim as an indication of the online integration of information from top-down predictions and bottom-up sensory evidence. Thus, with higher predictability of the incoming input, the MSPCstim shows a higher correlation between the input stimulus phase and the EEG signal. On the other hand, in comparison to MSPCstim, MSPCres is argued to be more strongly related to signal integration occurring at higher/later levels of the visual hierarchy as attention modulation had a greater effect on MSPCres. Specifically, when the stimulus is highly relevant/ attended, then the prediction error would be highly weighted and have a higher effect on expectations.

Thus, the HFT procedure provides a potential opportunity to explore the integration of bottom-up and top-down signals in the brain, whilst at the same time controlling for predictability and attention. However, the paradigm is still in its infancy and the division of the Multi Spectra Phase Coherence (MSPC) metric into stimulus-driven and response-driven has been used in only two studies so far without a control condition (Coll et al., 2020; Gordon, Tsuchiya, et al., 2019). Thus, further research is needed to determine its validity.

Here we want to further investigate validity of the paradigm using different stimuli and adding a control condition. We investigate the perception of biological motion through PLDs as it could allow us to determine whether differences in biological motion perception are due to reduced integration at a higher or at a lower level of the hierarchy. Furthermore, the paradigm could help us to understand whether attention plays a role in the high precision setting of prediction errors in ASD.

Although the original SWIFT tagging procedure utilised wavelet transformation (scrambling) of semantically meaningful images, in the present experiment, a global percept - a walking point-light display - is transformed by spatially scrambling the position of the dots that make the PLD. Due to the nature of the PLD stimuli (made up of dots), there would not have been enough visual information to use a wavelet transformation. Thus, we argue that oscillating the PLD between its intact and scrambled forms should produce a semantic tagging at one frequency, producing a SWIFT-like component observed in the electroencephalographic (EEG) signal. In this modification of the paradigm

attention is modulated by instructing the participants to count one of two alternating PLDs which have opposite walking direction. Predictability is modulated by adjusting the proportion of one direction to the other across blocks. To produce the low-level SSVEP component, the contrast of the dots (flickering), which comprise the PLD and the surrounding noise dots, changed at a second tagging frequency.

Beyond adapting the original HFT procedure to use a dynamic stimulus instead of static, we also aimed to use the autism quotient (AQ) questionnaire to observe how the SSVEP-component, the SWIFT-like component and the IMs vary with the change in autistic traits of the participants. Following HIPPEA's predictions, we would expect that in individuals with high AQ, the SWIFT SNR will decrease with the increase in predictability, but at lower predictability, it would be more easily disturbed by new events. Hence, the volatility of the environment will play an important part (Lawson et al., 2017; Palmer et al., 2017). Further, results by Coll et al. (2020) show that individuals with higher autistic traits had consistently higher SNR for the SSVEP, SWIFT and their intermodulation (IM) component signals. This was not investigated by the original authors, but it was observable in their graphical representation of the data. Following HIPPEA it would be expected that prediction errors should disturb predictions more in individuals with high AQ. At high predictability, where the counted stimulus is the one presented almost all the time, we would expect to see no difference between individuals with high AQ and low AQ. Thus, it would be interesting to observe the balance of the SWIFT and SSVEP signals. This is dependent on whether the effect of the SWIFT is observable as it was not replicated in Coll et al. (2020). Further, if ASD individuals put a lot more weight on prediction errors and bottom-up information, with increase of AQ scores we should see a reduction of the positive correlation (the slope) of the IM components signals with increasing predictability. In fact, Coll et al.'s (2020) findings indicate that with higher AQ the positive slope is reduced in comparison with participants with lower AQ.

Finally, if processing in autism happens according to HIPPEA, then we would expect to observe different MSPCres slopes with the change in AQ scores. If individuals with autism can make predictions, and those predictions are based on the narrow band of expectations they form, then the MSPCres slope, which is an

indication of the integration of the information at a higher level of the processing hierarchy, should be similar between individuals with low AQ and those with high AQ for the attended images. However, if at the same time autistic individuals put high weight on the precision of the prediction errors, then the slope should be either less negative, neutral or tending towards positive for the unattended images with the increase of repetition. Effectively, the higher precision setting of prediction errors would lead to very similar activation of attended and unattended images in the brain. This could be inferred from the results by Goris et al. (2018) where local deviants were not modulated by expectations but were equally detectable in the EEG signal. Similarly, MSPCstim, which is assumed to indicate the integration of information at earlier stages of the processing hierarchy, should be in the same direction as in the original studies, but since prediction errors would have a higher weighting, the slope of the relationship with the increase in predictability should also be shallower.

Due to the COVID-19 pandemic and the national lockdown in Scotland associated with it, the study was stopped prematurely. From hereafter, this chapter will use the collected data as a pilot and proof of concept for the feasibility of the study and the proposed paradigm.

5.3 Methods

5.3.1 Participants

Three participants were recruited to take part in the main experiment. Participants were students from the University of Dundee. Two participants were to be research assistants in the data recruitment and were not compensated, one participant was awarded experiment credits towards their first-year undergraduate Psychology degree. None of the participants suffered from neurological conditions or had a history of seizures. Participant characteristics are shown in Table 13 below.

Table 13. Participant characteristics

Participant	Age	Sex	AQ	Eyesight	Handedness
P1	21	Female	18	Corrected	Right
P2	21	Female	11	Normal	Left
P3	18	Female	10	Normal	Right

Since this is a new paradigm and the presentation time of the coherent biological motion figure is expected to be very short, two pilot studies were run to make sure that the participants can perceive the walking stimulus and its direction. These are described in Appendix E.

5.3.2 Measures and materials

5.3.2.1 Autism Spectrum Quotient (AQ)

The Autism-Spectrum Quotient has been originally developed for investigating the autistic traits in individuals. It has 50 items, divided into 5 subscales with 10 items each – communication, social skills, imagination, attention to detail and attention switching (Baron-Cohen et al., 2001). Participants were not excluded based on their scores of the AQ.

5.3.2.2 Paradigm characteristics

Point-light displays

The main paradigm consisted of a point-light display (PLD) that cycled from coherent to scrambled for 30 times within one trial. A PLD represents a human figure with light dots at the major joints. The PLD was presented at the centre of an LCD screen at a visual distance of ~ 75cm. The visual angle of the coherent walker spanned 3.43^{o} horizontally and 5.5^{o} vertically. This size was chosen as it is the most commonly used in the literature, and it was found that roughly this size provides the best signal to noise ratio in behavioural detection paradigms (Hiris, 2007). When the walkers were scrambled, the displays spanned 8.76^{o} horizontally and 9.24^{o} vertically. The larger visual angle of the scrambled display allowed for the scrambled PLD to blend in with the noise and to avoid crowding in the middle where the coherent PLD would appear.

The PLD was embedded in noise created by one left wrist joint and one right ankle joint from six scrambled copies of the PLD. This type of noise was chosen

as it maintained the same motion and energy as the original PLD (also see the 'Frames' pilot in Appendix E). The noise dots were also shifted in phase, to minimise distraction based on the simultaneous movement of the dots. Each of the six pairs was shifted by 20 frames starting from no shift to 100 frames after the presented coherent frame. Additionally, as there is a left walker and a right walker that swap which one is the coherent one, there were always two walkers present, in such an arrangement that when one of the walkers was coherent, the other one was scrambled. This added to the noise background. A sample video can be seen at https://figshare.com/s/d6ffd876ce684610349c.

To avoid EEG artefacts due to brain waves synchronising with the specific type of scrambling of the PLDs, three different scrambling versions were chosen for each of the two PLDs – left and right walkers. The scrambling variations changed at random for each cycle and the change of the scrambling variation happened at the point of the coherent walker, to avoid abrupt jumps of light dots, which could also create artefacts in the EEG signal.

In the control condition, the PLDs were inverted, in this way maintaining the same movement profile of the light dots. The control condition aimed to make sure that any tagging at the SWIFT-like frequency was not due to just the movement of the dots towards the centre of the screen but was due to recognition.

Frequency selection

We imposed several limitations when choosing our frequencies. Firstly, the frequencies need to be the product of the refresh rate divided by an integer: F1 – 60[4]/12 = 5Hz and F2 – 60/90 ~ 0.6667Hz for the SSVEP and SWIFT-like frequencies, respectively. Next, the fundamental frequencies, their harmonics and combinations, must not coincide with each other. This allowed for a better signal to noise ratio. Next, all frequencies – fundamentals, first harmonics and

[4] The final study utilised a monitor with a vertical refresh rate of 60Hz. The pilot studies utilised a CRT monitor at 120 Hz. The frequency calculations were equivalent for the 'dots' pilot and the final study: F1 – 120/24 = 5Hz and F2 – 120/180 ~ 0.6667Hz. The SWIFT frequency differed, due to the varying length of the presentation of the coherent biological motion.

their combination needed to be separated by at least 5 bins of the frequency spectrum for our chosen length of the trial – 45s and bin of 0.022ms.

SSVEP component. To create the SSVEP component of the paradigm, we chose a flicker at a frequency F1 = 5Hz. In this way, we were able to avoid the alpha-band 8-12Hz, and we chose a significantly high-frequency resolution to increase the SNR - a 512Hz sampling rate with vertical screen flip of 60Hz.

SWIFT-like component. Due to the nature of the paradigm - i.e., a moving stimulus, it is impossible to create perfect co/sine function presentation of the transition between scrambled and coherent walkers, while at the same time providing participants with enough information to be able to perceive the stimulus. For the representation of a movement, more than one frame is necessary. Thus, our presentation for the biological motion presents a sine function with a plateau. The frequency for the coherency modulation was chosen to be in the lower frequency spectrum – F2 = 0.6667Hz, as it better penetrates higher-level processing (Norcia et al., 2015). Moreover, the paradigms that have used SWIFT have shown that low frequencies provide a good SNR for processing at 1.2HZ, 1Hz and 0.8Hz (Gordon et al., 2017; Gordon, Hohwy, et al., 2019; Gordon, Tsuchiya, et al., 2019)

5.3.2.3 Paradigm

Participants were asked to count either the walker that appeared to walk to the left or to the right on each trial. Predictability (expectations) levels are categorised based on the proportion of right and left walkers appearing in each trial. This ranged from low predictability, when the left and right are shown 50% of the time, to high predictability, when one of them is shown 90% of the time. This variability was introduced, to modulate the predictability of the attended stimulus. Each trial contained 30 walkers and lasted ~45sec. Thirty-six such trials were created. The length of the experimental and the control trials were identical, with the experiment lasting ~1h.

Since the ability to perceive the biological motion is necessary to be able to say that we are accurately tagging semantic features (in this case the biological motion), we trained our participants to a 65% accuracy criterion. There were

two training blocks. In the first one, the coherent walker was coloured red only when it was coherent. After each trial (total of 6) participants were asked to say how many walkers they saw in the indicated direction. There were two possibilities at this stage – the proportions were either 50% or 80% for one of the directions. If participants performed at around 65% accuracy or above, then they moved to the second training block, where the coherent walker was not coloured. In other words, it was identical to the experimental paradigm. If on average they performed with 65% accuracy or above across both training blocks, they were included in the experiment.

5.3.3 Procedure

Participants were seated at ~ 75cm from the screen. They then filled in a brief form about demographics and the AQ questionnaire. During each trial, participants saw a flickering cloud of dots (at F1) and at specific time points (at F2) some of the dots formed the shape of a human figure – a point-light display. Participants were asked to mentally count the number of displays that show a person walking to the right or left. At the end of each trial, they were asked to manually input their response. The times of input were not restricted, and participants were encouraged to take a break to rest their eyes during this time, to avoid aftereffects. There were two blocks: the experimental block with the upright PLDs and a control block with the inverted PLDs. Before the beginning of the experiment, participants took part in the training procedure as described above. The EEG was fitted to the participants only if they passed the 65% training criterion. To maintain the training effects, all participants performed the experimental block first.

5.3.4 EEG data acquisition

Continuous EEG was acquired with a 64-channel ActiveTwo system (Biosemi). In addition to the EEG set-up, additional electrodes were placed around the eyes and on the face over facial muscles (measuring electromyography). Electrodes were placed according to the international 10/10 system. The vertical and horizontal electrooculogram were recorded by electrodes around the right eye and at the temples. An active and a passive electrode (CMS, and DRL,

respectively) were used as a reference. Data were sampled at 512Hz for all participants.

5.3.5 Analysis procedure

5.3.5.1 EEG data analysis.

Data pre-processing was performed using EEGLAB toolbox in MATLAB. Each block/condition was pre-processed separately. The pre-processing steps do not differ between the two conditions.

Participant data pre-processing. Data were initially referenced to the central channel – Cz, high pass filtered at 0.4Hz, the European power frequency and its harmonics were removed [50 & 100 Hz], and the mean of the channels was removed. Next noisy channels were identified through visual inspection and were replaced with spherical spline interpolation. Initial data cleaning was performed through independent components analysis and the number of components was reduced to 30 using principal components. Components representing blink and muscle artefacts were identified and removed – average removed components: 15.3 (SD = 4.76; range: 10 – 23). The large number of components being removed was caused by participant 2 having a very noisy data as evidenced in Table 14. After the identified components were removed, data were re-referenced to the average of all channels.

Next, automated data cleaning was performed based on the procedures reported by Coll et al. (2020) and Gordon et al. (2017; 2019). Cycles of the biological motion were considered noisy when 2% or more of the sample points were noisy. If a cycle was termed noisy, the cycle was replaced on each channel with the average signal for that trial. If more than 10% of the cycles in a channel were considered noisy, the channel was replaced with spherical spline interpolation. Individual trials were excluded if more than 10% of cycles in a trial were termed noisy after interpolation. Due to the small number of participants, none were excluded based on number of excluded trials (reported in the Results section).

Spectral analysis. This analysis was done only on the electrodes under consideration described below. Firstly, the last few milliseconds of the trials were removed to leave trials of equal length - ~44.998s. Power-spectra was

extracted at the tagging frequencies - F1 and F2, their first harmonics - 2F1 and 2F2, and for their second order IMs – F1 + F2 and F2 - F1. This was achieved by applying the Fast Fourier Transform (FFT) over the predefined period of the trial with zero padding to 45ms to allow for a frequency resolution of ~0.022Hz. We implemented a multitaper frequency transformation with a single Hanning taper. Frequency band of interests were set between 0.3Hz and 23Hz. Signal-to-noise ratio (SNR) was calculated by dividing the amplitude at each frequency by the mean amplitude of 10 neighbouring frequencies (five on each side).

Channels under consideration.

In previous studies, the FFT and the SNR ratio calculation was performed over all of the electrodes in one of the studies (Gordon, Tsuchiya, et al., 2019) and on a set of 30 of the 64 electrodes in the following studies (Coll et al., 2020; Gordon et al., 2017) - centroparietal, temporo-parietal, parietal, parieto-occipital and occipital electrodes CPz, CP1, CP2, CP3, CP4, CP5, CP6, Pz, P1, P2, P3, P4, P5, P6, P7, P8, POz, PO3, PO4, PO5, PO6, PO7, PO8, Oz, O1, O2, T7, T8, TP7, TP8. Thus, these electrodes were used for the analysis.

Further, biological motion perception has been localised to the parietal and temporal regions. Specifically, a network has been identified that involves F7, P3 and Pz (Fraiman et al., 2014). These electrodes were able to differentiate between biological motion and scrambled motion and clustered together in a functional network. Hence, we also looked at F7 as it showed the fastest response to the biological motion in Fraiman et al.'s study (2014). Further, since biological motion observation is observed in the supplementary motor system, electrodes Cz, C3 and C4 were also included (Ulloa & Pineda, 2007). These electrodes have also been examined in autism as examples of the mirror neuron system (Todorova et al., 2019). As the SSVEP is supposed to invoke lower-level visual activation, we expect the occipital electrodes to capture SSVEP signal, whereas SWIFT has been originally observed over central electrodes, and any effects would therefore be more prominent there (Koenig-Robert & VanRullen, 2013).

5.3.5.2 Analysis

In the original analysis plan, an LME model with expectation, attention, AQ and the combinations of two-way and three-way interactions was going to be included as fixed effects. Random effects were going to include a random intercept for frequency nested within channels nested within participants and random expectation, attention and interaction slopes for each participant. However, since we only have three participants, results are interpreted descriptively.

The performance was only considered for trials in which participants showed more than 50% accuracy. This cut off affected mostly P3's data, where the SSVEP and SWIFT signals were not artificially pulled down/up by bad performance. In the control condition, since the walkers were inverted, the accuracy of the participants was meaningless as the assumption was that the inversion would interfere with the perception of a coherent walker. Nevertheless, P1 shows a higher accuracy than the other two participants. This indicates that they might have been able to detect the inverted movement of the PLD or used motion cues to infer direction.

The predictability factor was calculated based on the left walking direction of the PLD – i.e., if 90% of the stimuli were right walkers, then 10% were left walkers. The attention factor was determined in a similar manner – if participants were instructed to count the left walker, then stimuli were attended, if they were instructed to count the right walker, they were unattended.

MSPCstim and MSPCres analysis for the 2nd order IM components was performed on a within-trial level using an adapted code provided by Noam Gordon. Each trial was trimmed to 45s and was divided into 15s epochs, with 1s steps. Phase angle differences were calculated between the phase of the IM components (4.333 and 5.6667) and the phase of the input frequencies (SWIFT and SSVEP), to extract MSPCstim. Phase angle differences were also calculated between the IM components' phase and the phase of the FFT output for the input frequencies to extract MSPCres.

For the MSPC analysis, following the procedure in Gordon, Tsuchiya et al. (2019) and Coll et al. (2020), we calculated a new attention factor. The factor recoded the trials into attended trials - when the walking direction that was counted (attended) was the one presented more than 60% of the time, and unattended trials, where the walking direction that was counted (attended), was not the one that was presented more than 60% of the time. The trials when the left/right walking distribution was 50% were excluded.

As only a descriptive analysis is performed here, all graphs represent the average scores for the averaged regions of interest.

5.4 Results and Discussion

EEG data quality for all participants for each condition is presented in Table 14. Overall, data loss was minimal, except for P2 in the control condition, where most of the trials were lost.

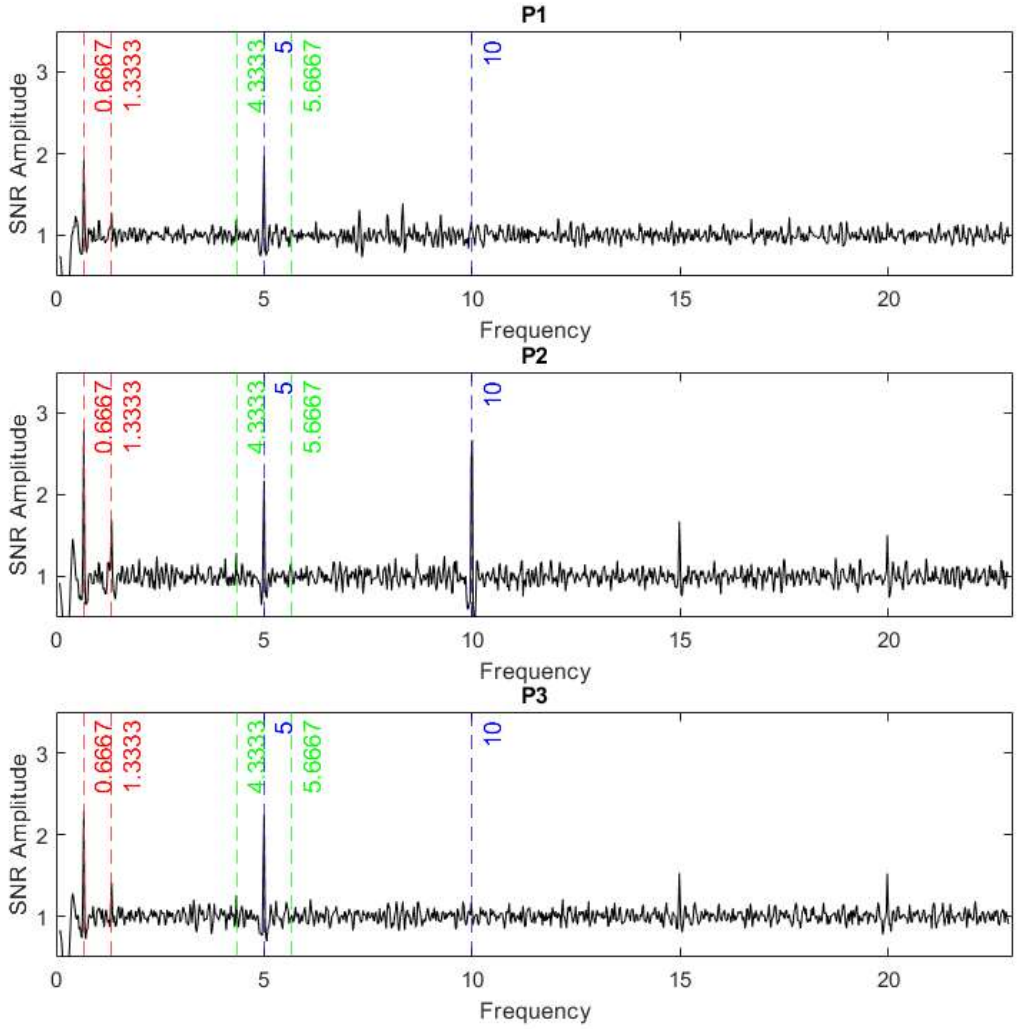
Table 14. Participant EEG quality characteristics

Participant	Experimental condition			Control condition		
	Noisy cycle (%)	Noisy channel (N)	Noisy trials (N)	Noisy cycle (%)	Noisy channel (N)	Noisy trials (N)
P1	0.65	5	0	1.89	4	3
P2	4.54	5	6	10.65	5	21
P3	3.13	2	0	3.13	2	1

5.4.1 Brain activity entrainment

After cleaning we verified that the frequency-tagging procedure was able to entrain the EEG brain activity. Figure 16 shows the result from the FFT averaged over the channels of interest and all the trials for each participant. Peaks can be seen at the two tagging frequencies – F1 = 5Hz (SSVEP) and F2 = 0.6667Hz (SWIFT). A distinguishable peak is also observed at the SWIFT harmonic at 2F2 = 1.3333Hz. A peak at the SSVEP harmonic at 10Hz was only observable in P2. The IMs, however, are not distinguishable from the surrounding frequencies. For the IMs, only the SSVEP-SWIFT(4.3333Hz) combination shows a peak, however, that peak is not clearly distinguishable from the surrounding frequencies. For completeness, the observations will be reported for the IMs as the small peak at SSVEP-SWIFT, could become more pronounced in a larger sample, where the noise between participants is averaged out. Nevertheless, these observations need to be viewed with extreme caution.

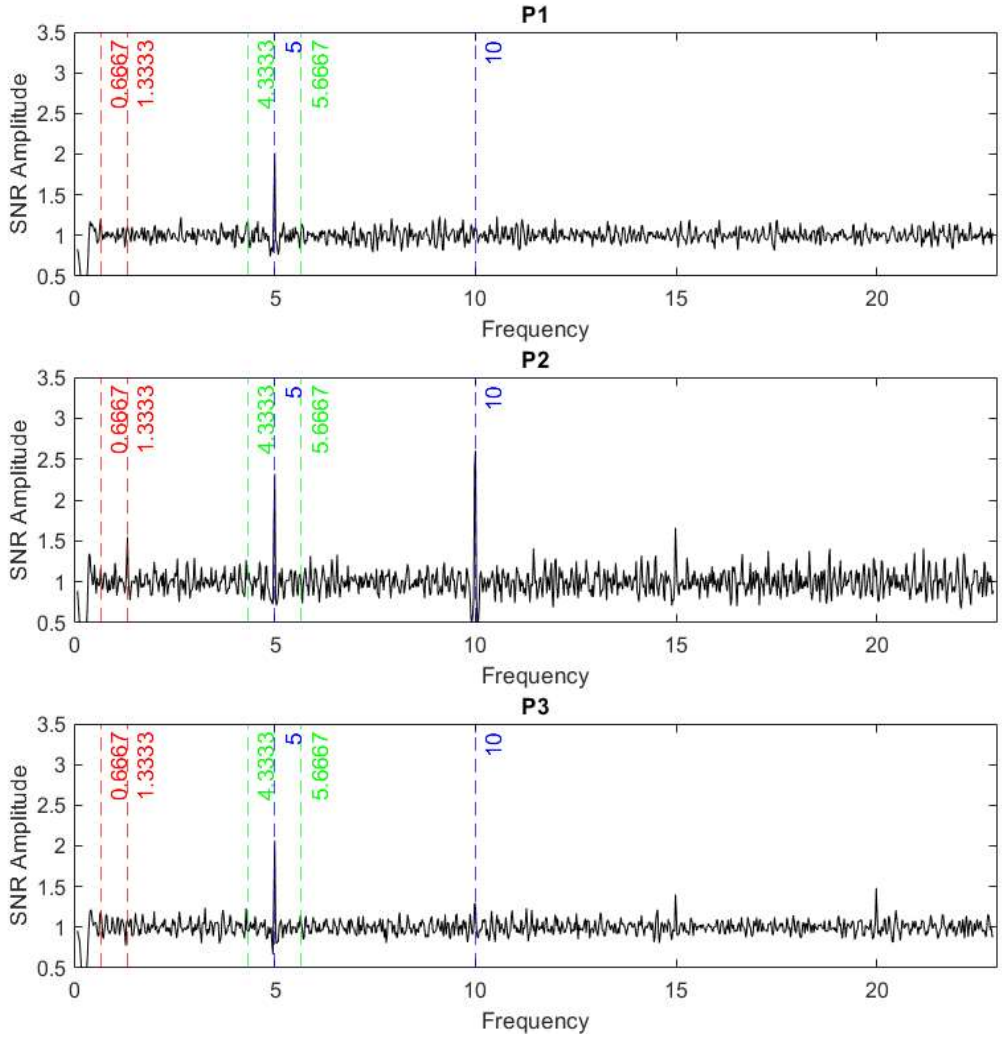
Figure 16. Frequency spectrum of the experimental condition for the SNR Amplitude for each participant.



Note: **the SWIFT and its harmonic are presented in red, the SSVEP and its harmonic are shown in blue; the two IMs are shown in green.**

In the control condition the SSVEP is clearly observable at 5Hz, however we cannot see any distinguishable peaks at the SWIFT tagging frequency – F2 = 0.6667Hz. Again, only P2 shows a strong SSVEP harmonic at 10Hz.

Figure 17. Frequency spectrum of the control condition for the SNR amplitude for each participant.



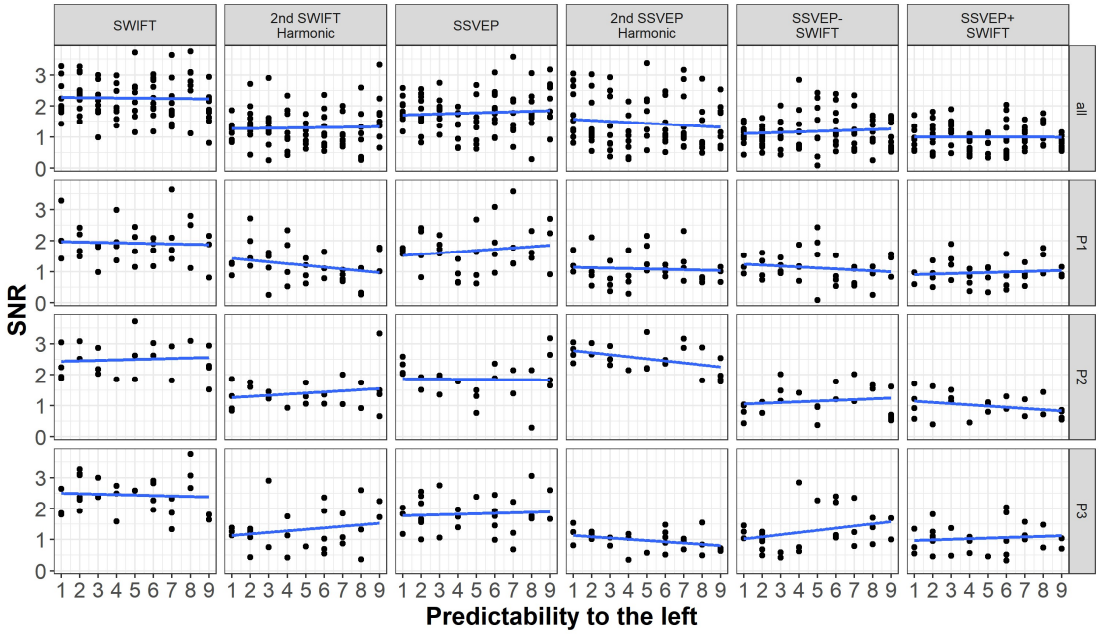
Note: **The location of the SWIFT frequency and its harmonic are marked in red, the SSVEP and its harmonic are shown in blue; the two IMs are shown in green.**

5.4.2 Experimental condition

As seen in Figure 18, the SNR for the SWIFT frequency does not show a trend in either direction in the combined plot. This is probably due to the opposing trends between participants with a shallow decrease in the signal for P1 and P3 and a shallow increase for P2. For the SWIFT harmonic, P1 shows a decrease in the SNR as predictability increases, whereas P2 and P3 show an increase. The SNR for the SSVEP frequency is showing an increase with predictability for P1 and P3, however, this trend is stronger in P1 than P3. P2 on the other hand, does not show a distinguishable pattern in the SSVEP with an increase in predictability. The SSVEP harmonic was not evident in P1 and P3 in the spectral analysis (see Figure 16). This is also evident in the size of the SNR for these

participants in comparison to P2. Nevertheless, an increase in the SNR with an increase in predictability is evident in 2/3 participants to various degrees. Further, looking at the SNR plots, there is a stronger pattern in one of the IMs (SWIFT - SSVEP) where P2 and P3 show an increase in SNR as opposed to P1. In the second IM (SWIFT + SSVEP), P1 and P3 show an increase in the SNR. It is noteworthy that in the spectral analysis only the SWIFT – SSVEP IM produced a peak at the expected frequency, although the SNR was not distinguishable from surrounding variability, thus the effects in the IM should be interpreted with caution.

Figure 18. SNR change with predictability level for each participant and their combination.



These results correspond to the original findings by Gordon et al. (2017) and Coll et al. (2020) to a small degree. The shallow increase in the SNR with the increase in predictability in the SSVEP signal was also evident in the figures from both papers (Coll et al., 2020; Gordon et al., 2017). In relation to the SWIFT, two of the participants show a small decrease in the SNR for the SWIFT-like signal and two of the participants show an increase in one of the IMs. Similarly, we see a decrease in the SWIFT signal in the original paper by Gordon et al. (2017), but not in Coll et al. (2020). Thus, it could be said that the SWIFT signal is in line with the original findings to an extent. The SWIFT harmonic, however, shows a slight increase in the SNR with an increase in predictability. This is opposite to previous findings. These discrepancies could be one of the reasons

why Coll et al. (2020) was not able to find an effect at the SWIFT frequency – i.e., a cancellation of the effects between the original frequency and its harmonic. Moreover, the slopes are shallower than those in the SSVEP, hence, the effects might be very small and sensitive to the stimuli used/participant variability.

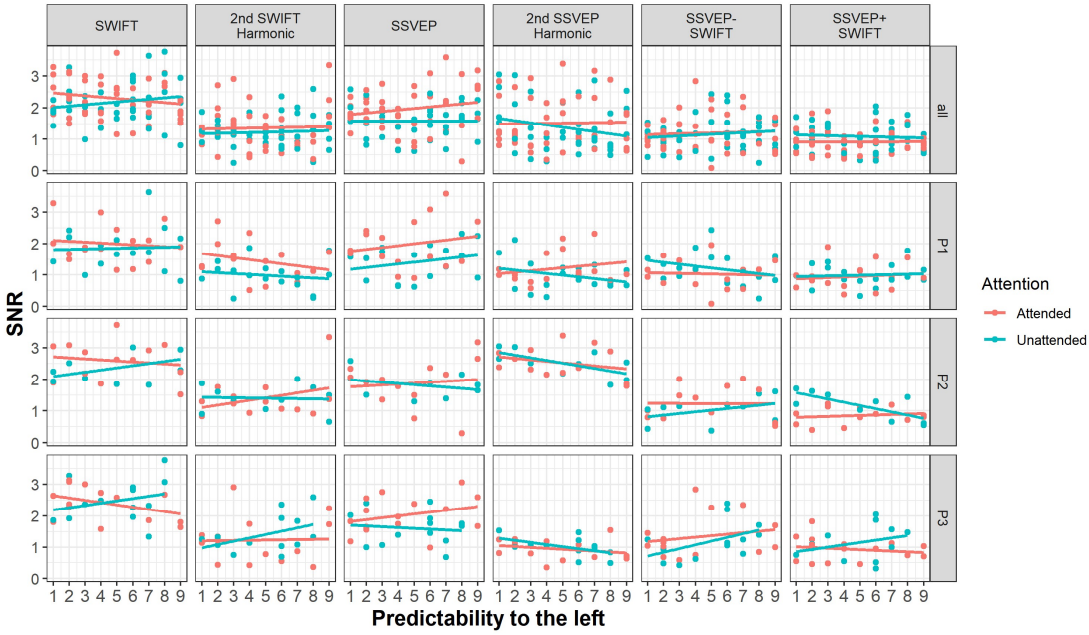
From the point of view of the IMs, a peak was found at the predicted combined driving frequency SSVEP-SWIFT, but not SSVEP+SWIFT. Its overall pattern of an increase in the SNR with an increase in predictability follows the findings from Gordon et al. (2017) and Gordon, Tsuchiya, et al. (2019) and Coll et al. (2020). This suggests that with the increase in predictability there was an increase in the integration of the top-down and bottom-up signal (Gordon et al., 2017). It is noteworthy that the previous studies looked at the 2nd and 3rd order IMs and we only looked at the 2nd order components, as no consistent peaks were observed between participants in the rest. Furthermore, we also did not observe a strong signal for the IMs, especially for the SSVEP+SWIFT IM. We also saw substantial variability between the three participants. Thus, these effects are more likely to be evident within a larger sample.

5.4.2.1 Effect of attention and predictability

We also wanted to see how the effects of attention affect the SNR. Thus, the trials were separated into attended and unattended with respect to the leftward walking PLD. In Figure 19**Error! Reference source not found.**, we can see that attention plays a small part in differentiating any effects. The effect at the SWIFT frequency seems to suggest that attention leads to a decrease of the SWIFT signal, whereas inattention leads to an increase in the SNR with an increase in predictability. The SNR at the SWIFT harmonic shows an increase regardless of attention level. This difference between the two frequency graphs appears to be driven by one very high value by P2. If we consider the SWIFT signal as an indicator of top-down processing (Gordon et al., 2017), then the decrease in the SNR for attended stimuli with an increase in predictability would be indicative of better processing, and reduced activity for stimuli that fit with the prediction. On the other hand, the increase in the SNR for unattended stimuli could be indicative of an increase in top-down control as the individual

must actively disregard the expectation of the more often repeating stimulus to perform the task at the attended stimulus.

Figure 19. SNR change with predictability and attention level for each participant and their combination



With respect to the SSVEP, the SNR is larger at the attended level than at the unattended. This is driven by participants P1 and P3. The SSVEP harmonic shows a different pattern in all three participants. However, the signal strength was only easily detectable in P2 (see Figure 16). It is possible that the extra processes evoked by the counting could have inadvertently contributed to the larger SNR in the attended images. However, as the SSVEP is argued to reflect bottom-up processing, which would not depend on higher process (Gordon et al., 2017; Gordon, Tsuchiya, et al., 2019), neither expectation nor attention should affect the signal as the flicker of the dots is not dependent on either.

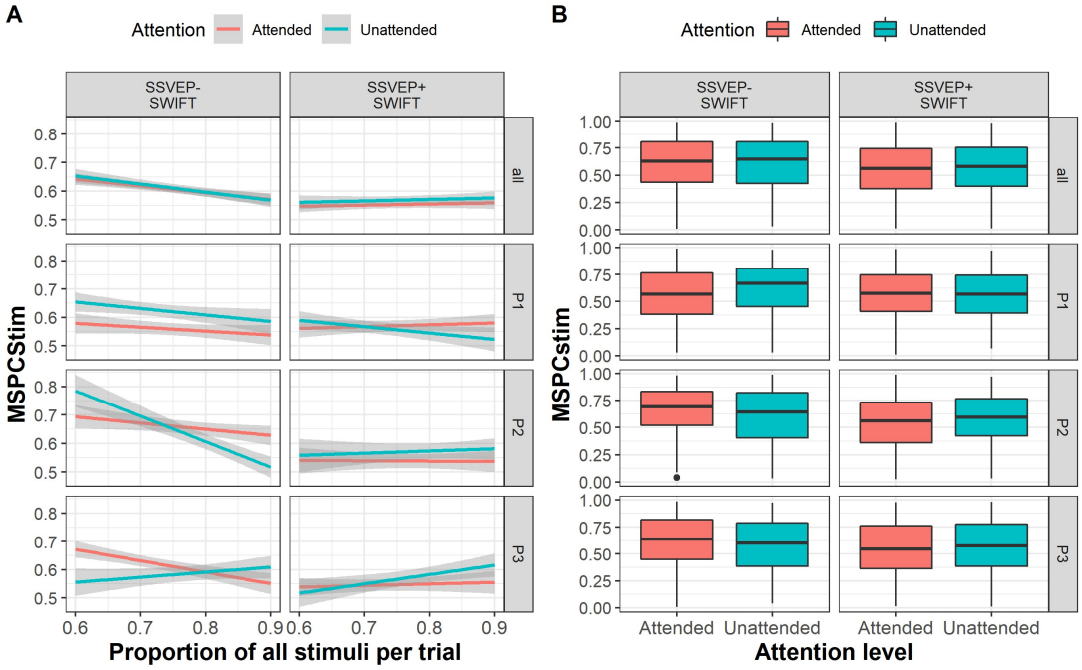
Finally, observing the interaction between expectation and attention at one of the IM components (SSVEP-SWIFT), an increase in the SNR with an increase in predictability was present regardless of attention. For the SSVEP + SWIFT IM, this was only present in P2. Further, the SNR of the IMs is much lower in comparison to the major frequencies and their harmonics. However, the size of the SNR for the IMs was also substantially smaller in the original studies by Gordon et al. (2017) and Coll et al. (2020). Thus, any effects would be difficult to distinguish within a small number of participants. The original research did not explore the

effects of attention on the SNR; therefore, we cannot make a direct comparison here.

5.4.2.2 MSPC analysis

Moving to the MSPC analysis, we see that the low SNR of the IMs leads to some contradictory findings. As seen in Figure 20 the MSPCstim did not show an effect of attention. The results for the MSPCstim of SSVEP - SWIFT IM for P1 and P2 attended and unattended trials showed a decrease in MSRCstim with increase of proportion of stimuli per trial and P3 showed an increase for unattended and decrease for attended trials. All three participants also showed a decrease in MSPCstim with increase in proportion of stimuli per trial. Thus, overall results appear to show the opposite effect in comparison to Gordon, Tsuchiya et al.'s (2019) and Coll et al.'s (2020) findings. Whereas their data shows an increase in the MSPCStim with repetition, the present results show a decrease in MSPCStim. The results for the MSPCstim of SSVEP + SWIFT IM however appear to show results similar to the results by Gordon, Tsuchiya et al. (2019) and Coll et al. (2020). Despite this, since the SSVEP + SWIFT frequency was not distinguishable from the surrounding frequencies in the spectral analysis shown in Figure 16, any interpretation of this result should be done with caution.

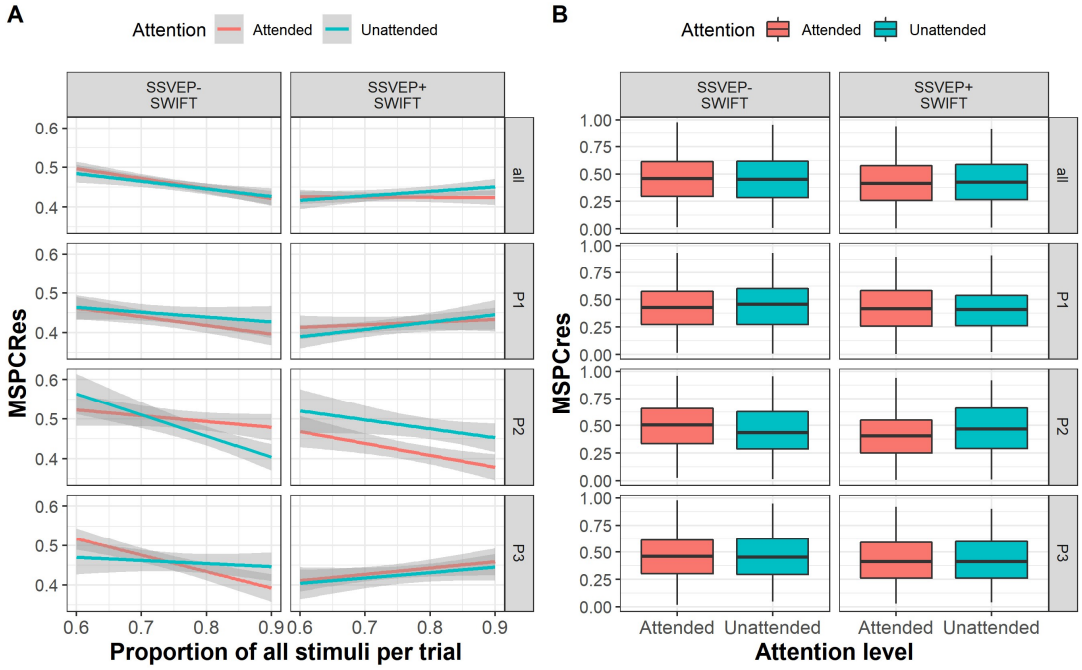
Figure 20. MSPCstim for each 2nd order IM for each participant and overall.



***Note*: A: Change in MSPCstim values with increase of the proportion of all stimuli in a trial each attention level (0.6 – 60% of PLDs in the trial; for the claculation refer back to Section 5.3.5.2); B: Difference in MSPCstim between the two attention levels.**

Similarly, the MSPCres, also does not appear to show a large distinction between attended and unattended trials (see Figure 21). In fact, the results mirror those from the MSPCstim. The only difference appears to be that the MSPCres is overall lower than MSPCstim. Further, the effects appear opposite to the ones found in both Gordon, Tsuchiya et al. (2019)and Coll et al. (2020). In all participants, attended trials were more likely to lead to a decrease in spectral coherence with the increase of the proportion of attended stimuli per trial. The same was true for unattended trials, however with a less steep slope. The MSPCres for the SSVEP+SWIFT IM showed an increase for unattended trials.

Figure 21. MSPCres for each 2nd order IM for each participant and overall.



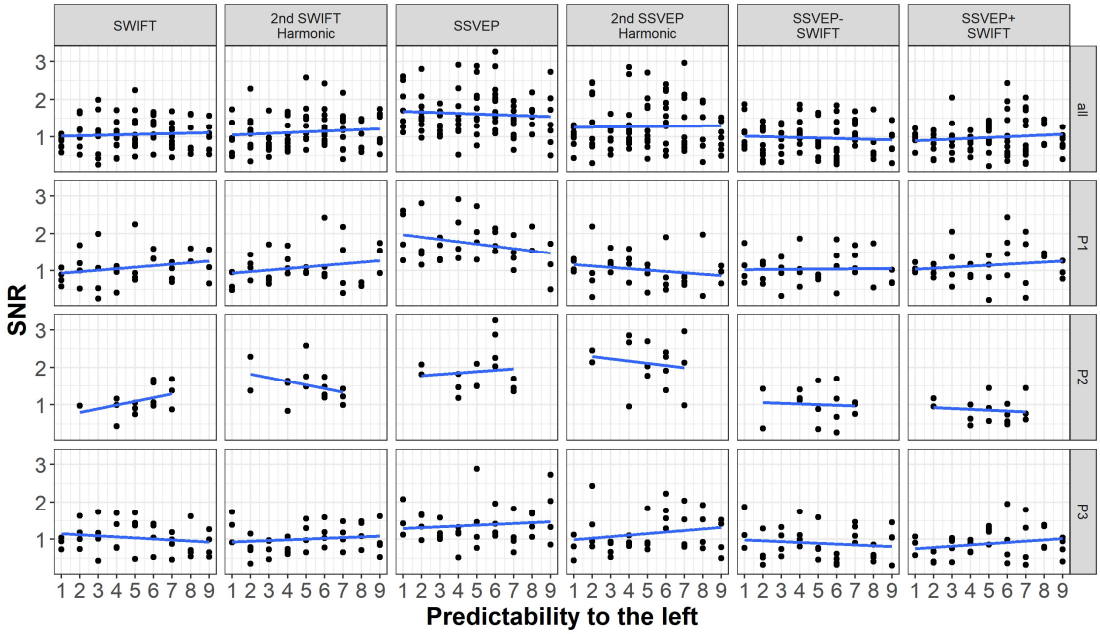
***Note*: A: Change in MSPCres values with increase of the proportion of all stimuli in a trial each attention level (0.6 – 60% of PLDs in the trial; for the claculation refer back to Section 5.3.5.2); B: Difference in MSPCres between the two attention levels.**

5.4.3 Control condition

In the control condition, the signal appears flat for most of the observed frequencies (see Figure 22). As P2 had a large number of removed trials in this condition, their results are difficult to interpret. The SWIFT driven SNR, as well as the SWIFT harmonic, appear to show an increase with predictability. This could be due to the fact that P1 showed a high accuracy within the control condition, although there was not a visible indication of high SNR for the SWIFT frequency (see Figure 17). Nevertheless, the SWIFT signal is lower than that in the experimental condition. Importantly, the SSVEP which is only connected to the flicker of the stimuli shows a large SNR, despite the opposing effects in P1 and P2. However, the overall effect appears to lead to a decrease of SNR with predictability, driven by P1. On the other hand, P3 still shows an increase with predictability. As in the experimental condition, only P2 showed a large SNR for the second SSVEP harmonic, which is evident in the lower SNR for the other participants. The SNR for the IM components also appears flat, except the SSVEP+SWIFT IM, which shows a shallow increase in SNR with an increase in predictability. As expected, neither of these frequencies appeared on the frequency spectrum as seen in Figure 17. However, these small effects in the

linear combinations suggest that the modulation of predictability can influence the overall signal, not only at the tagging frequencies.

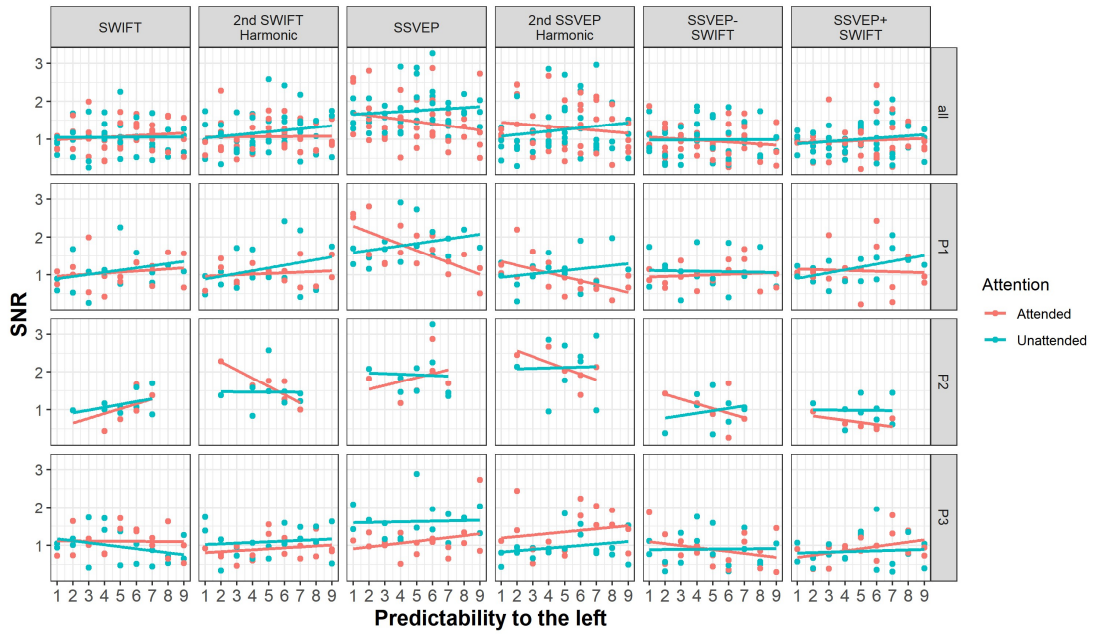
Figure 22. SNR change of predictability level.



5.4.3.1 Effect of attention and predictability

The additional separation by attention levels in Figure 23 does not provide any further insight into the performance. This is expected since the stimulus was not supposed to be interpretable. The most distinguishable differences are seen in the SSVEP SNR. Unlike the experimental condition, attended stimuli led to lower overall SNR, except in P3 for the SSVEP harmonic. The reduction in the SNR with increase in predictability for the attended trials in P1 is interesting, as they showed high accuracy in the control condition, despite the inverted stimulus, and the direction with increase in predictability is in the opposite direction from the experimental condition. Thus, it is possible that attention does not play a highlighting role and increase the SSVEP SNR as we suggested in the previous section.

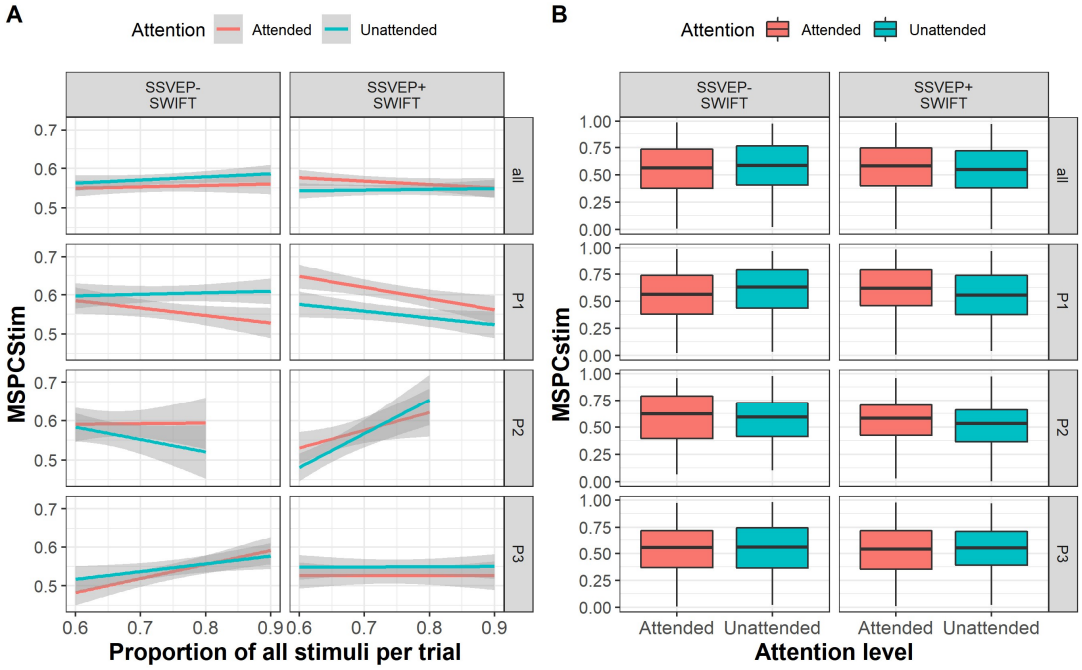
Figure 23. SNR change with predictability and attention level for participant 1 and 3 and their combination



5.4.3.2 MSPC analysis

If the MSPC components reflect the integration of information at the different levels of the hierarchy, we should not observe any effects, in the control condition. This, however, does not appear to be the case. Looking at the MSPCstim in Figure 24, the effects do not appear to differ much from the ones in the experimental condition, although, the overall MSPCstim is lower. The main difference is seen in P3, where attended stimuli lead to an increase in phase coherence with the increase in the proportion of attended stimuli per trial, for the SSVEP-SWIFT. This was in the opposite direction for the experimental condition.

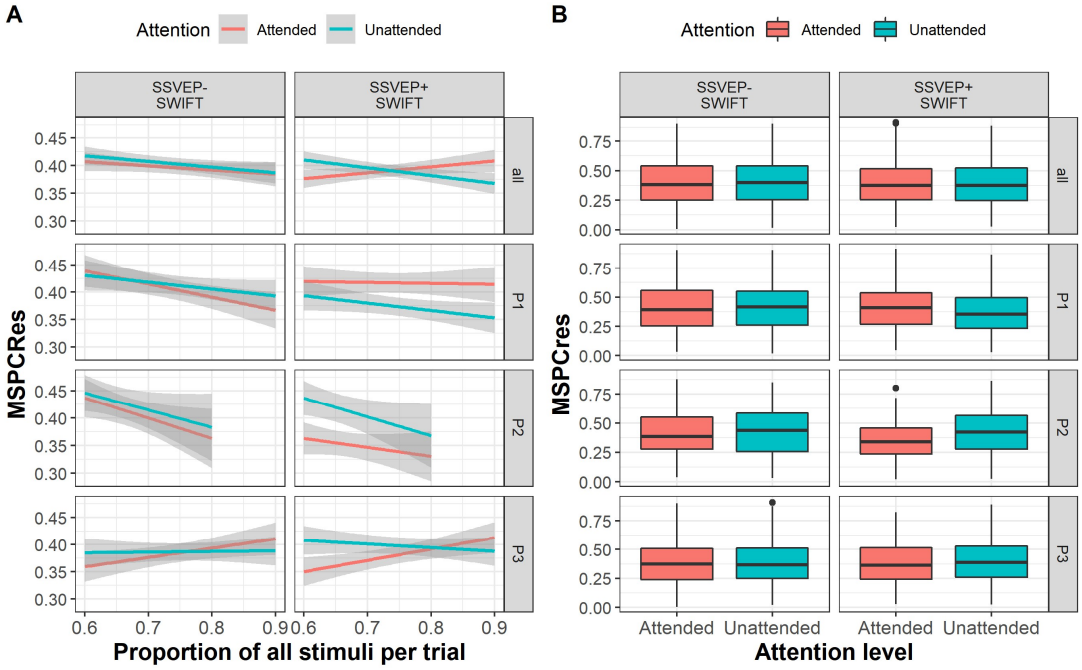
Figure 24. MSPCstim for the 2nd order IMs for each participant and overall.



***Note*: A: Change in MSPCstim values with increase of the proportion of all stimuli in a trial each attention level (0.6 – 60% of PLDs in the trial; for the claculation refer back to Section 5.3.5.2); B: Difference in MSPCstim between the two attention levels.**

With respect to MSPCres, similar to the experimental condition, the overall strength of MSPCres is lower than that for MSPCstim (see Figure 25). MSPCres for the SSVEP-SWIFT frequency shows a similar but less negative relationship between MSPCres strength and predictability for both attention conditions in comparison to the experimental condition. Additionally, unlike in the experimental condition, P3 shows an increase in MSPCres with increase or predictability for attended trial. For the SSVEP+SWIFT IM, there is an overall increase in the MSPCres for attended trials with the increase in the proportion of attended stimuli and a decrease for unattended trials. This is however largely driven by the P3. In comparison, in the experimental condition we observed a general increase for both attention conditions in P1 and P3.

Figure 25. MSPCres for the 2nd order IMs for each participant and overall.



***Note*: A: Change in MSPCres values with increase of the proportion of all stimuli in a trial each attention level (0.6 – 60% of PLDs in the trial; for the claculation refer back to Section 5.3.5.2); B: Difference in MSPCres between the two attention levels.**

5.5 Discussion

In this proof-of-concept experiment, we aimed to determine 1) whether the effects that are observed in the original studies would be replicated, , 2) whether the effects are specific to the modulation of the paradigm by the inclusion of a control condition and 3) whether the HFT paradigm could be applied to dynamic stimuli like a PLD walker.

As evident by the SNR results for predictability and the interaction with attention, we see that the effects observed in earlier papers are present to similar degrees in our data. Examining the interaction with attention, we can also see that attention overall increased the SNR for the SWIFT and SSVEEP frequencies and their harmonics. However, the difference is much smaller between the attention levels for the IM components. Nevertheless, these effects are certainly encouraging for the use of dynamic stimuli in an HFT paradigm. However, the absence of distinguishable IM components is worrying. This is especially important considering the controversial findings from the MSPC analysis in comparison to the observed trends in previous studies and could indicate a problem with the present paradigm, a difference due to the dynamic

nature of the stimuli or a lack of specificity of the MSPCstim and the MSPCres. As the effects of the SNR analysis appear to follow the earlier findings, it is unlikely that the paradigm itself is the cause of this observation. Thus, further research into the validity of MSPC analysis is necessary.

Moving on to the MSPC analysis, as a reminder, the MSPC metrics aim to represent the integration of bottom-up and top-down information at a lower-level (MSPCstim) and a higher-level (MSPCres) of the processing hierarchy. In our results, we observe opposing patterns in the participants that do not coincide with the original findings. For MSPCres and MSPCstim, in the experimental condition all participants show a decrease in coherence with increase in predictability for the attended trials which is opposite to the patterns observed in Coll et al. (2020) and Gordon, Tsuchiya et al. (2019). For unattended trials, there was also a reduction in MSPCres with an increase in the proportion of attended stimuli , which was in accordance with the earlier findings. Thus, the lack of substantial overlap in the experimental condition and the original research findings from Gordon, Tsuchiya et al. (2019) leave the use of MSPC as an open question.

So far, we have discussed the effects observed in the experimental condition. However, the mark of a well-working paradigm is its specificity and previous research has not included a control condition. Thus, it was necessary to investigate if the effects are due to the modulation of attention and prediction, or whether the effects would be observed even in the event of no global stimulus to drive the SWIFT frequency. When examining the effect of predictability in the control condition, the SNR is very weak for the SWIFT and the IMs and mostly flat. The strength of the SSVEP, however, stands out. Thus, in the absence of a global, contextual modulation, the SSVEP, which is driven only by the luminance properties of the dots is still present in the signal. This is an indication that the SSVEP is indeed capturing low-level information. The effects of attention are difficult to interpret in a meaningful manner. Although P1, shows some interaction between predictability and attention, this could be related to their ability to detect the direction of walking of the inverted stimulus. However, the minimal differences in the rest of the conditions and participants are expected. This is because the task – i.e., count a specific PLD walking direction, was not meaningful under the assumption that they were not

able to detect the PLD. So far, the experimental modulation appears to show good specificity.

Looking at the MSPC results, P3 shows an increase in MSPCres with an increase in predictability of the attended targets. Although this is in-line with the original findings, the fact that the consistency comes from the control condition where the SWIFT was not detected, indicates, that the MSPC analysis may not be specific to the existence of the signal integration at the levels of the hierarchy suggested by Gordon, Hohwy et al. (2019). Similar to the experimental condition, for unattended trials, there was also a reduction in MSPCres with an increase in the proportion of attended stimuli. It is important to note that both the MSPCstim and the MSPCres are generally weaker in the control condition in comparison to the experimental condition. Since Gordon, Hohwy et al. (2019), implicated MSPCstim as an indication of the online integration of information from top-down predictions and bottom-up sensory evidence, whereas the MSPCres as an indication of integration at higher levels of the hierarchy, this finding suggests that since there are no IM components in the control condition, any MSPC effects will also not be visible to the same extent as when the IM is present, as there is not integration occurring.

Our findings are interesting for two reasons. Firstly, they show that with the absence of a contextually interpreted global stimulus which the SWIFT is tagging, the SNR values for the IMs and the strength of the MSPCres and MSPCstim are reduced. Thus, this validates the role of the SWIFT modulation of the signals with the use of a dynamic stimulus, and further for the static stimuli part of Gordon et al. (2017; 2019) and Coll et al. (2020). Secondly, these findings suggest that the MSPCres and MSPCstim tend to show effects over the IMs despite the absence of an effect in the IMs themselves in the control condition. This brings into question the specificity of the two measures. If there are no effects in the IMs, then we should not expect the MSPC components to indicate effects in the control condition. Moreover, the absence of overlap between our and the earlier published findings in the MSPC analysis further suggests that the effects might be specific to static stimuli, or to the use of faces/houses. Thus, the MSPC measure may not be able to extend to the overarching interpretation of representing the effects of bottom-up and top-down integration in the brain as suggested by Gordon, Hohwy et al. (2019).

These interpretations come with several caveats. Firstly, we have only three participants, on whom we are basing these conclusions, and the dataset for one was very noisy. It is possible, that these participants would be outliers in a larger sample. Therefore, these results need to be further expanded upon with a larger sample. Secondly, the use of a dynamic, as opposed to static stimuli could have introduced an additional component into the task. For example, the present task was more difficult than the tasks used by Gordon et al. (2017; 2019) and Coll et al. (2020), with participants needing to incorporate several moving targets into a whole and separate them from the moving noise in the background. Thus, task demands might be masking some of the effects, particularly in the MSPC measure. Furthermore, the original paradigm utilises a sine wave whereas we utilised a wave with a plateau. This was driven by the nature of the stimuli - a dynamic stimulus would require more than one frame to be interpreted. Thus, the IMs that we chose to look at, might not represent the most appropriate combination of the two tagging frequencies. However, when observing the frequency spectrum graphs of all frequencies no other prominent frequencies are present. In our 'Frames' pilot we discovered that individuals were better at detecting biological motion with fewer frames, thus using fewer frames might be a reasonable step forward. However, participants' performance was very variable and thus, a reliable choice of frames would be difficult to make without a larger sample.

Nevertheless, the HFT paradigm gives the potential opportunity to investigate how the integration of bottom-up and top-down information occurs in individuals who have been diagnosed with autism and those with high amounts of autistic traits. This would allow finding evidence to provide support for or against the HIPPEA theory and other predictive coding models of autism. The recent findings by Coll et al. (2020) are an indication of the potential explanatory power that this type of paradigm could have in elucidating the effects of autism. They found that the MSPCres strength was influenced by attention, expectation and autistic traits. Whereas differences were not seen for the attended images, a significant effect was seen for unattended images. Specifically, individuals with higher autistic traits showed a negative slope in the unattended images whereas the rest of the participants showed a positive one. Surprisingly, however, the individuals with low autistic traits showed a positive MSPCres slope with

increased predictability for unattended images. These MSPCres findings are surprising as the individuals with low AQ seem to show the opposite effects to the two experiments in the original work by Gordon et al. (2019), where typical individuals show a negative slope indicating reduced integration at higher levels for unattended images. These results are difficult to interpret, especially in light of the non-replicated effect of the SWIFT in the study by Coll et al. (2020).

The differences in findings between the current study and within previous research, necessitate further investigation to elucidate if the same results could be found in experiments with different paradigms, such as the one presented here. However, before continuing forward, further work needs to be done to clarify the specificity of the MSPCstim and MSPCres and their relationship to the SWIFT and the SSVEP and the proposed underlying mental processes. For example, predictability is confounded with repetition, as with higher predictability inevitably one of the stimuli would be repeated several times in a row. This could then lead to effects of repetition suppression in the brain which might obscure the effects of predictability. Thus, it might be necessary to consider an alternative modulation of predictability.

The HFT paradigm is in its infancy and the division of the MSPC measure into stimulus-driven and response-driven has been used in only two studies so far with slightly contradictory findings (Coll et al., 2020; Gordon, Tsuchiya, et al., 2019). The novelty and the potential explanatory power of a paradigm that claims to separate the effects of attention and expectation and to investigate the integration of bottom-up and top-down information necessitates extensive research to determine its validity. Moving forward, our results and the results by Coll et al. (2020) need to be confirmed and replicated to provide a clearer picture of the present observations. Most importantly, however, is our recommendation for future studies to include control conditions to allow them to make less equivocal conclusions about the results they are observing.

Overall, the HFT paradigm has good potential for observing the effects of attention and prediction in the brain. Moreover, here we show that using a dynamic stimulus is possible and that it can achieve similar results in the main parts of the paradigm. Thus, this type of paradigm could be used in future work to allow the investigation of the combination of bottom-up and top-down signals

in a dynamic context with the investigation of autistic traits and autistic individuals.

Chapter 6 General Discussion

This thesis investigated whether the main component of the HIPPEA theory (Van de Cruys et al., 2014, 2017) – namely, the invariably high precision setting of prediction errors - is a characteristic seen in autistic individuals' performance in motion and biological motion tasks. We investigated this phenomenon under different attentional conditions, since attention is considered to be the mechanism through which precision is estimated (Feldman & Friston, 2010). In this thesis, we used different experimental set-ups to directly assess the interplay between attention and response of autistic individuals or individuals with high autistic traits with respect to prediction errors as defined by unexpected events.

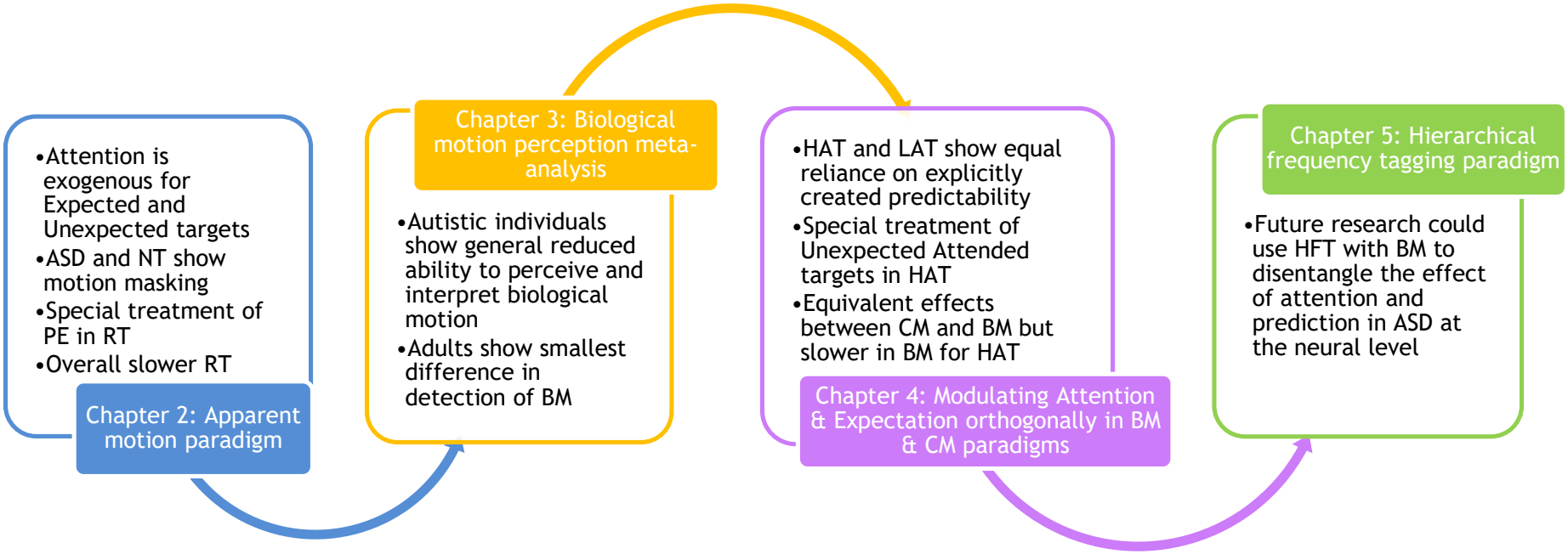
The findings of this thesis build upon previous and ongoing research investigating the predictive coding perspective of autism and provide some clarity to the effects of attention and expectation. More broadly, we show that both autistic individuals and individuals with autistic traits can form expectations about the regularity of events. This is also true when the regularity is explicitly stated. Moreover, this ability does not appear to be affected by the use of social stimuli such as biological motion. However, the task is more difficult as would be expected from previous research in biological motion perception in autism. Additionally, we see that there is a special treatment of prediction errors in autistic individuals and those with higher autistic traits. These findings can help move research forward, by providing insight into the balance of the weighting of prediction and prediction error in autism. Specifically, it appears that although the use and development of predictions is not affected, prediction errors still receive special treatment. Thus, to be able to disentangle these effects, this thesis further provides a proof-of-concept study for an EEG paradigm that could allow the investigation of predictions and prediction errors in the brain and their interplay.

6.1 Summary of findings

This section will present a summary of the findings of the discussed chapters and will show how each of the findings feed into each other to provide a coherent

understanding of the treatment of prediction errors in autism. A summary of findings can be seen in the flowchart presented in Figure 26 below.

Figure 26. Schematic representation of the findings in this thesis.



Note: **PE – Prediction Error; BM – Biological motion, CM – Coherent Motion; HAT/LAT – High/Low Autistic Traits; HFT- Hierarchical Frequency Tagging**

In Chapter 2, we used an apparent motion paradigm to test the treatment of prediction errors in autistic individuals. As has been shown previously, neurotypical individuals tend to experience the flashing squares part of the paradigm as a single moving square (Alink et al., 2010; Sanders et al., 2012; Schwiedrzik et al., 2007). This percept is assumed to represent a top-down model of the brain. Thus, a stimulus that appears on the apparent motion trace, with its position consistent with the predicted position of the 'moving' square, would be more easily perceived because of the expectation of a moving stimulus in that location. On the other hand, if a stimulus appears outside of those parameters, it would be considered unpredictable – a prediction error - and it will be missed more often. We observed that both the NT and autistic participants detected fewer targets when the presentation timings of the top and bottom flashing stimuli conformed to an apparent motion presentation. At the same time, in both groups more participants detected a larger proportion of predictable than unpredictable targets. This further supports the hypothesis that autistic individuals are able to form and use predictions/top-down models (Van de Cruys et al., 2014) and is in opposition to theories suggesting the formation of weak or uninformative priors (Pellicano & Burr, 2012). We observe, however, that more ASD participants showed faster detection of unpredicted targets than NT participants. This falls under the hypothesised higher weighting of prediction errors (Lawson et al., 2014; Palmer et al., 2017; Van de Cruys et al., 2014), as unpredictable events appear to exogenously attract attention to a larger extent. Although we were not able to recruit the full sample to allow us to say with certainty if prediction errors are treated differently, these findings are indicative of a difference in the way the autistic brain is treating unexpected events.

As Chapter 2 suggests, the autistic brain appears to treat prediction errors differently, although this does not appear to be to the detriment of the established prediction. It is possible that motion perception is not 'stressing' the system enough for any effects to be visible and autistic individuals might be utilising alternative strategies to produce the same behavioural effects as neurotypicals. One way to increase the complexity of the task, without complicating the requirements for the participants is to use more complex stimuli such as biological motion. At the same time, we need to be sure that any

increase in complexity is not going to obscure the effects of interest. Thus, in Chapter 3 we conducted a meta-analysis looking at biological motion perception in autistic individuals (Todorova et al., 2019). We investigated different types of tasks such as biological motion detection, action perception and understanding, and emotion detection from point-light and full-light stimuli. We found that there was a consistent difference between NT and autistic groups in the accuracy of detection/interpretation of biological motion stimuli across all tasks. However, these differences were smaller in the simple detection and action perception tasks, whereas the emotion perception tasks proved to be the most difficult. Moreover, we observed an effect of age – with the increase in age, the difference between autistic and neurotypical individuals decreased. These findings then dictated the use of simple biological motion, whose interpretation would be less ambiguous in Chapter 4 and Chapter 5 – namely, a PLD kicker and a PLD walker, respectively.

Carrying on from Chapter 2, and informed by the results of Chapter 3, in Chapter 4 we used a modified Posner paradigm where we independently modulated predictability and attention based on the paradigm used by Kok, Rahnev et al. (2012). We explicitly instructed participants that in some blocks a target will appear on the left or right side of the screen 75% of the time, whereas in other blocks we told them that there is no expectation about where the target is going to appear. From Chapter 3 we know that simple biological motion stimuli should be easily perceived by autistic individuals despite still presenting difficulties. Thus, the attentional cue that we used was a PLD of a human performing a kicking motion. Chapter 4 also presented results from a second experiment, where the attentional cue was a circle of directionally moving dots, instead of the PLD. In this way, if the social component of the cue was obscuring any effects, we would be able to observe differences in the tasks. We also decided to recruit participants from the general population and use their autism quotient scores as continuous predictors and a proxy for the effects we would see in autistic individuals. This was done because this is a new paradigm set up, which would warrant first investigation in a neurotypical population. This also allowed us to increase participant numbers from online recruitment and avoid the lengthy period of recruitment experiences in Chapter 2.

Similar to the autistic individuals in Chapter 2, we saw that individuals with high AQ scores were able to use the statistical regularities of the environment – i.e., they were able to establish a predictive model that was beneficial for their performance. Additionally, as HIPPEA argued, we observed differences, albeit non-significant, in the processing of unpredictable cues specifically in the attended condition for individuals with high AQ scores. These results were similar to the descriptive findings in Chapter 2. The fact that this was true only for the attended, but not the unattended unexpected targets puts into question the inflexibility of the HIPPEA theory, but it is consistent with arguments about reduced attentional capacity in autism (Orekhova & Stroganova, 2014). We also descriptively observed weaker effects of attention with increase in AQ, which was evident when no expectation was set (None condition). This suggests that when the environment was volatile, higher AQ scores led to larger reliance on incoming sensory information. However, when an expectation was set, individuals with higher AQ did not disregard it and instead used it to produce similar performance gains as their lower AQ counterparts. Since we only looked at individuals with different AQ scores, these effects then could be more pronounced and/or qualitatively different in individuals who have been given an ASD diagnosis. This chapter emphasised the importance of investigating attention and predictability in the same paradigm when trying to understand the relevance of the predictive coding perspective in autism.

The findings from Chapter 4 appear to contradict the 'inflexibly high precision' of prediction errors in autism as suggested by HIPPEA (Van de Cruys et al., 2014), however they do correspond to an extent with the arguments from Palmer et al. (2017), who suggest that precision setting is context dependent in autism and could be related to the precision associated with volatile environments as argued by the aberrant precision model (Lawson et al., 2014). Thus, it appears that the interaction between the prediction error and the prediction is where the differences in the processing styles between autistic and non-autistic individuals lie. However, this difference would be difficult to investigate at a behavioural level and could be more evident in differences in the neural processing between the two groups. In fact, differences in processing networks that produce similar behavioural performance have been previously observed in autistic individuals (eg., McKay et al., 2012).

Following this, in Chapter 5 we proposed the use of a hierarchical frequency tagging (HFT) paradigm, which modulates attention and predictability, and investigates the interaction between the prediction error and the prediction at both a lower and higher-level of the neurocognitive hierarchy. We provide a proof-of-concept study of a modification of Gordon et al.'s (2019) task, where we used a PLD walker presented in noise that oscillated between a coherent and a scrambled state. We modulated predictability by varying the number of times the walker was walking to the left or right direction and we also modulated attention by asking participants to count only one of the walking directions. Although we only recruited three participants due to COVID-19 restrictions, we were able to show that this modification can elicit tagging of the low-level information – i.e., the dots forming the PLD - and the higher-level information- i.e., the coherent walker. We were also able to show that the effect of the higher-level tagging is specific to the participants being able to consciously perceive the walker, rather than an artefact of the oscillation of the stimuli. This type of paradigm could provide an opportunity to investigate their interaction in autism. However, the results from these participants do not allow us to fully replicate the effects of interaction between the lower-level and higher-level information by looking at the intermodulation components, which were found in previous studies (Gordon et al., 2017; Gordon, Tsuchiya, et al., 2019). Thus, we hope that the development of this paradigm and its initial small test will serve as a guide to future research when trying to disentangle the effects of both attention and predictability in autism and when trying to characterise the differences in the interaction of prediction errors and the prediction between neurotypical and autistic individuals.

Taken together, these studies help to improve our understanding of predictive processing in autism and lead to a clear conclusion and future directions. The precision-weighing in autism is aberrant in respect to prediction errors. However, this does not impede the use or the establishment of predictions in the form of statistical regularities in the environment. Instead, it could be the case that it is not a trade-off between the precision setting of the prediction error and the prediction, rather it is the new events that are particularly salient in contrast to the prediction. Moreover, attention appears to play a key role. However, this is only the case when expectation is already at play. Using an

experimental design like the HFT proposed in Chapter 5 could allow the disentanglement of where the differences occur between autistic and neurotypical individuals.

6.2 Implications for predictive coding theories in autism

As stated above, precision setting of prediction errors in autism appears to be set differently compared to neurotypical individuals. This is in accordance with the HIPPEA theory and its main component of high of prediction errors. It appears, however, that the findings are unable to provide support for the inflexible part of the HIPPEA theory. This has been echoed by both Lawson et al. (2017) and Palmer et al. (2017). Learning at every instance caused by inflexibly high precision- of prediction errors will slow down the ability of individuals to use already established models. We see this slowing down in the reaction times in the BM experiment in Chapter 4 and in published research (eg., Lawson et al., 2017). However, if the precision setting was inflexibly higher, we should observe faster or more accurate responses for unpredictable events regardless of attention levels. We see this descriptively in Chapter 2 where attention is exogenously captured by the targets themselves, as well as descriptively in Chapter 4, where attention is directed endogenously, and the targets themselves capture attention exogenously. However, in Chapter 4 we do not see faster responses for targets that are unexpected but also unattended. Hence, when both attention and prediction are in favour of a particular event, the prediction errors are weighted higher, but not when attention is allocated elsewhere. To further establish this link, rigorous controlling of attention and prediction needs to be adopted when researching the predictive coding perspective of autism.

It could be argued that the link between attention and predictability is not entirely clear due to the descriptively observed IOR effect in the CM experiment presented in Chapter 4. However, we also observe that the effects of attention are not as large with the increase in AQ. This suggests that with the incorporation of expectation, which individuals with autism and high autistic traits are clearly able to rely on, attention provides an additional guiding mechanism that assists in dealing with information in the environment. Thus, the processing of attentional cues would be slowed down, minimising any IOR effect

for this population. This could lead to the attentional cue having a facilitatory effect when there is extra cognitive load added by the processing of the prediction.

If matching the input to the prediction in autism is based on low-level information (Van de Cruys et al., 2014), this process will be more effortful than if the matching is done on an abstract level. The different effect cognitive load has in autistic in comparison to neurotypical performance has also been observed in eye-tracking and neuroimaging studies (Skripkauskaite et al., 2021; Wadhera & Kakkar, 2020). Hence it is possible that although predictions assist in behavioural performance, they increase the cognitive load for autistic individuals. It has also been shown that, in the presence of high perceptual load, perceptual sensitivity is reduced in NT individuals, but this reduction is smaller in autistic individuals (G. Murphy et al., 2016; Remington et al., 2012). Taking these two points together, it is possible that since autistic individuals are better at detecting information under larger perceptual load, the cognitive load added by the more effortful matching of incoming information to the prediction makes them more likely to detect unpredictable targets, but only when they are facilitated through attention. This could explain the effects of faster detection of unpredictable attended targets in both the CM and the BM experiments in Chapter 4. Thus, we see higher weighting of prediction errors, although this is only true in specific contexts, which might be dependent on the interplay between perceptual and cognitive load introduced in the paradigm (G. Murphy et al., 2016; Palmer et al., 2017).

The discussion above, however, is speculative and it is unlikely to be resolved by adding more complexity to the tasks to increase the cognitive load for certain experimental paradigms. Instead, observing the interaction between prediction errors and predictions on a neural level is necessary. Paradigms and set-ups like the one proposed in Chapter 5 could allow us to develop a deeper understanding of what effects prediction and attention have on the autistic cognitive and perceptual processing pathways. More importantly, the proposed HFT paradigm does not rely on the coupling effect of attention and predictability. This is because attention is not driven by the task, and instead by the instructions to participants to count a certain stimulus. This could potentially go around the

questions asked in Chapter 4 of whether attention or prediction drives perceptions and behavioural responses.

Additionally, paradigms that investigate the effects of attention and expectation at the auditory level could provide further insight into the interaction of prediction errors and the prediction. One example is the paradigm presented by Chennu et al. (2013) where participants were presented with auditory sequences with a) global standard with two deviant options - type varying sound in same ear, or same sound but in opposite ear and b) global standard containing a type variant with two deviant options - type varying sound in same ear that matched the standard sequence, or same type variant as in the global standard, but in opposite ear. In this task participants were asked to either count the rare sequences, the deviant tones, or to perform a demanding visual task. The authors then investigated the early activated MMN component, the later expectation modulated P300 and the contingent negative variation component in EEG measurement, which the authors argue indicates the consolidation of salient stimuli in consciousness. This paradigm, like the HFT paradigm discussed earlier, utilises attentional demands and predictive contingencies to investigate the interaction of the two factors in the brain. Hopefully, such paradigms will allow researchers to develop a better understanding of the interplay between attention and predictability in autism and by extension further test the assumption of high precision of prediction errors in autism.

The predictive coding perspectives of autism provide a new avenue for researchers to attempt to understand autistic individuals' experience. However, these theories are still in their infancy (see review by Cannon et al., 2021) and the role off attention is still not clear. Thus, it is important that research continues to allow for these theories to evolve.

6.3 Limitations and future directions

The interpretations in this thesis come with several methodological and more general caveats. Firstly, sample sizes are small in this thesis. Chapter 2 provides a well-controlled paradigm, however, the study is not sufficiently powered. This is not to say that the descriptive findings are not without their merit as they present an effect that is also observed in Chapter 4 in respect to the detection

of unpredictable events. This makes the findings in Chapter 2 worth replicating with a larger sample. However, large samples of individuals diagnosed with autism are difficult to obtain, especially for smaller studies like the psychophysical ones reported here. In fact, most research in autism uses quite small samples (Simmons & Todorova, 2018; Todorova et al., 2019), which is due to the difficulty in recruiting participants, and finding matching controls. To remedy this, groups of several research labs doing the same study and pooling their small sample sizes could improve our understanding of the larger effects in autism and provide more rigorous results. As a first step, we are sharing the analysis and paradigm set up code, along with the majority of the raw data for the studies in Chapter 2 and Chapter 4 with the idea that other researchers will could collect more data to add to our understanding of precision setting in autism.

Secondly, the proposed HFT paradigm in Chapter 5 contains only three participants, whose results could be outliers in a larger sample. This is further indicated by the complete opposite patterns found in the measures of MSPCstim and MSPCres – i.e., the two measures that would indicate the interaction of the low- and higher-level information in the brain. Nevertheless, we were able to replicate the low-level SSVEP and high-level SWIFT tagging in all participants, and there were indications of intermodulation components, which due to the small number of participants were too noisy to allow for firm conclusions. Intermodulation components usually show the smallest SNR (Gordon et al., 2017; Gordon, Tsuchiya, et al., 2019), therefore in larger samples where individual noise is cancelled out to an extent, these could be more prominent.

Thirdly, in the experimental part of this PhD thesis, participants were always given the task to respond to a target. These active tasks will inevitably tunnel attention to ‘achieving’ the required goal (Van de Cruys et al., 2014). If autistic individuals are more sensitive to prediction errors, then being provided with a task that requires their attention but still leaves enough resources for external stimuli to break through would be the way to understand to what extent there is high weighting of prediction errors when the prediction errors are not something that individuals are actively searching for. In this way, a less relevant task, that is not passive but is also not goal-focused, alongside a less demanding goal-focused task would be one way to investigate these phenomena. This is an

important distinction from a completely passive task, where there is no task given to the participant (eg., Goris et al., 2018). As discussed in Section 1.2.4.2, completely passive tasks show differences that are eliminated in active tasks (Dunn et al., 2008; Keehn et al., 2013), which could be caused by differences in spontaneous allocation of attention (Orekhova & Stroganova, 2014). Due to these potentially confounding factors, effects in fully passive tasks might be difficult to interpret.

It is, however, difficult to study spontaneous attentional shifts without alerting the participant to the task. Potential candidate tasks are neuroimaging and eye-tracking experiments, where the instructions follow the original set up in Kok, Rahnev et al. (2012). In their experiment participants were only instructed to respond to targets appearing on one side of the screen. Thus, looking at the neural activation in the brain to the different categories of stimuli (attended/unattended – as driven by the central cue, and predicted/unpredicted as driven by the block wise prediction), researchers could in theory observe what is the difference in the detection of attended and unattended unpredictable targets in autism and whether the differences lie only in the processing of events that fall under those attention/prediction categorisations.

Fourth, as with all research on autism, the sample characteristics need to be considered (Simmons & Todorova, 2018; Todorova et al., 2019). In Chapter 2 we only recruited autistic males. Although this was done with the attempt to minimise heterogeneity, it is important for these results to be replicated and extended to female individuals. It is well known that there are not only differences in presentation of autism between males and females, but there are also differences in the their underlying neuroanatomy, which could lead to differences in perception and interaction with the environment (Antezana et al., 2019; Cummings et al., 2020; Floris et al., 2021; Smith et al., 2019). Hence it would be unreasonable and a disservice to the autistic community if observations in males are directly extrapolated to females.

In contrast, Chapter 4 used both male and female individuals. However, we relied on participants from the general population, and we used autistic traits as a proxy. Although autistic traits show a large spread in the general population (Ruzich et al., 2015) and can have a reasonable sensitivity and specificity in

respect to autism diagnoses (Brugha et al., 2020; Conner et al., 2019), the effects observed in these individuals will inevitably be smaller than those effects observed in autistic individuals. Thus, although we observe similar effects between the autistic participants in Chapter 2 and the participants with high autistic traits in Chapter 4, we should not extrapolate to autistic individuals directly, and instead the results should be replicated and extended in autistic individuals.

6.4 Conclusions

The results presented in this thesis contribute several important findings in relation to the predictive coding theories of autism. Firstly, we found that autistic individuals are able to form predictions and are able to use explicitly stated statistical regularities of the environment to guide their behaviour. In this respect, our findings are not consistent with arguments of weak priors in autism (Pellicano & Burr, 2012). Secondly, we observe consistent findings of increased precision setting of prediction errors as defined by unexpected targets in our tasks. This is consistent with arguments about high precision of prediction errors as suggested by HIPPEA (Van de Cruys et al., 2014) and by the aberrant precision model (Lawson et al., 2014).

However, despite HIPPEA arguing for inflexible precision setting, we do not see an inflexible high weighting of prediction errors, as unattended unpredictable targets in Chapter 4 were not detected faster by individuals with higher AQ scores. Thus, this part of the HIPPEA theory could not be corroborated. In addition, we also cannot corroborate Lawson et al.'s (2017) proposition about overestimation of volatility, which would lead to the higher precision of prediction errors. Although we speculate a larger reliance on sensory information in a volatile setting, the high precision setting of prediction errors is also evident in the stable environment (when expectation is set) in Chapter 4. To an extent, the findings from this thesis corroborate the contextual dependence of precision setting discussed by Palmer et al. (2017).

Although this thesis had the goal to test whether precision to prediction errors is invariably high in autism, our findings did not fully support this hypothesis. However, the findings corroborated that in autism higher precision is given to

prediction errors in certain contexts, dependent on the task. Although these results need to be replicated in larger samples, the findings should be taken as evidence against the ‘I’ in the HIPPEA theory. Thus, it is worth reviewing the theory in terms of recent findings from this thesis and the general scientific community and to incorporate discussions from Palmer et al. (2017).

Appendices

Appendix A

Below are the search strategies used for the extraction of the papers form the five electronic databases.

MEDLINE. An example search strategy for MEDLINE® (1946 to November week 1 2017 (OVID) is shown below:

((autis* or asd or asperger* or "childhood schizophrenia" or kanner* or PDD-NOS or PDD* or "pervasive development* disorder*") and (PLD* or "biological motion" or "human motion" or "point-light display*" or "action observation*" or "action observation network*" or AON)).tw. limit 1 to english language

WEB OF SCIENCE. In Web of Science the following search strategy was used:

TS = ((autis* or asd or asperger* or "childhood schizophrenia" or kanner* or PDD-NOS or PDD* or "pervasive development* disorder*") AND (PLD* or "biological motion" or "human motion" or "point-light display*" or "action observation*" or "action observation network*" or AON))

Timespan: All years.

Search language=English

PsycINFO. The search strategy for PsycINFO (EBSCOhost) is shown below:

TX (autis* or asd or asperger* or "childhood schizophrenia" or kanner* or PDD-NOS or PDD* or "pervasive development* disorder*") AND TX (PLD* or "biological motion" or "human motion" or "point-light display*" or “action observation*” or “action observation network*” or AON) TX (autis* or asd or asperger* or "childhood schizophrenia" or kanner) and (PLD* or "biological motion" or "human motion" or "point-light display*" or “action observation*” or “action observation network*” or AON) Limiters – English Search modes – Boolean/Phrase

Dissertations & Theses A&I (ProQuest) and Dissertations & Theses: UK & Ireland (ProQuest). The search strategy for Dissertations & Theses A&I (ProQuest) and Dissertations & Theses: UK & Ireland (ProQuest) is shown below.

all((autis* OR asd OR asperger* OR "childhood schizophrenia" OR kanner* OR "pervasive development* disorder*" OR PDD-NOS OR PDD*)) AND
all(("biological motion" OR "human motion" OR PLD* OR "point-light display*" OR "action observation*" OR "action observation network*" OR AON)) AND
la.exact("English")

The number of extracted records from each database can be seen in the Table 15 below in descending order.

Table 15. Number of records extracted from each database.

Database	Records returned Search 1	Records returned Search 2
Web of Science	483	102
PsycINFO (EBSCOhost)	163	1
MEDLINE® (OVID)	115	19
Dissertations & Theses A&I (ProQuest)	22	2
Dissertations & Theses: UK & Ireland (ProQuest)	10	NA

Appendix B

The strength of the body of evidence was assessed on both the study and review level using the Weight of Evidence approach developed by the EPPI-Centre (Gough, 2007) following the guidelines of Popay et al. (2006). On the study level this includes the following four criteria on which each study will be judged (see Table 16 below):

- Weight of evidence A (WOA)– Trustworthiness [taken from the quality assessment score, score obtained from quality assessment score ranges from 0 – 1: 0 – 0.333 scored as low, 0.334 – 0.666 scored as medium, 0.667 – 1 scored as high]; For fMRI papers an assessment was done using relevant criteria from the Standard Quality Assessment. Specifically, questions related to analysis and results were excluded but the fMRI methodology was assessed for robustness. This was done collaboratively by the authors.

- Weight of Evidence B (WOB) – appropriateness of the studies' research design in terms of the current research question. This reflects the quality of the eligibility criteria. If the eligibility criteria were specific enough, then all studies should contribute to the interpretation of the body of evidence;

- Weight of Evidence C (WOC) – This refers to the focus of the studies and whether their findings are generalizable to the question in hand;

- Weight of Evidence D (WOD) – This refers to the overall score for each study, which is sum of the other three components. This is the score that is given to each study in Table 16 and Table 17 in the main manuscript referred to as WoE (Weight of evidence).

Results for the overall weight of evidence and the sub-sections are shown below in Table 17.

Table 16. Weightings for Weight of Evidence C

Criteria	Weightings	Rationale
Sample	3. (high) Fully describes sample (e.g., Diagnosis criteria (ADOS, ADI-R, clinical diagnosis, 3Di), Gender ratio, Age (mean, SD), FSIQ/VIQ/PIQ or other intelligence measures (mean and SD), Presence or absence of additional diagnosis was specified (TD and ASD), specifies the characteristics that TD individuals were matched to ASD individuals) 2. (medium) Misses or not fully specifies one or two of the above listed sample characteristics (excluding diagnosis criteria as this was part of the eligibility criteria) 1. (low) Misses more than two of the above listed sample characteristics	This will allow to determine how generalizable the findings are to a wider population of individuals on the autism spectrum
Task	3. (high) Fully describes the paradigm used, the procedure and the participants' task during the experiment; the procedure is randomised/counterbalanced 2. (medium) Elements about the paradigm are missing or participant's task is not clearly stated. 1. (low) The paradigm is poorly described; participants' task is not specified	This will allow us to judge specific characteristics of the task and whether they could have contributed to any of the findings

Table 17. Weight of evidence average scores and standard deviation (SD) for each element.

	WoA	WoB	WoC	WoD
Average	2.963	3	2.454	8.417
SD	0.191	0	0.354	0.398

Appendix C

An exploratory analysis was conducted on EEG studies to investigate the effect of the used frequency or measure to evaluate differences between ASD and NT individuals. The analysis showed that there was an overall effect of the type of frequency or measure used – $F(3,25) = 3.1922$, $p = 0.0411397$. If the frequency that was used was between 11-13 Hz, the estimated effect was very large – $g = 2.452$ [SE = 0.80788054, 95% CIs: 0.7883932 – 4.115708] $t(25) = 3.04$, $p = 0.00554$], the effects were much smaller, but still large if the frequency was 8-13Hz $g = 0.8652822$ [SE = 0.38671, 95% CIs: 0.0702849 – 1.6603794] $t(25) = 2.248$, $p < 0.034115$]. If the frequency was 8-10Hz, the effect went in the opposite direction, with individuals with ASD showing larger but non-significant mu-suppression ($g = -0.4031$ [SE = 0.5201185, 95% CIs: -1.474211 – 0.6648] $t(25) = -0.78$, $p = 0.445542$]). Finally, if N100 was measured, the effects were very small and non-significant ($g = 0.2361$ [SE = 0.9907878, 95% CIs: -1.80447984 – 2.27605] $t(25) = -0.24$, $p = 0.81361$]). It is noteworthy that only one study looked at frequency 8-10HZ10Hz, one study looked at 811-103Hz and one study looked at N100. Thus, these results are unreliable, which can be seen from the large standard errors and confidence intervals. No further analyses were performed.

Appendix D

Biological motion

Student sample

In Table 18 are reported the data for the students. Variations between different conditions are small for accuracy indicating that individuals were able to perform the task. Additionally, reaction time appears faster for Expected targets in comparison to Unexpected targets. Further, Attended targets elicited shorter reaction times in comparison to unattended target.

Table 18 Participant accuracy and median reaction time

Attention congruency	Prediction congruency	Mean accuracy (SD)	Median RT (MAD)*
Attended	Expected	0.916 (0.277)	0.515 (0.120)
	None	0.897 (0.304)	0.529 (0.127)
	Unexpected	0.861 (0.346)	0.534 (0.124)
Unattended	Expected	0.909 (0.288)	0.534 (0.125)
	None	0.890 (0.313)	0.533 (0.125)
	Unexpected	0.887 (0.317)	0.563 (0.145)

***Note*: * MAD = Median absolute deviation.**

The overall results from the models are consistent with the overall effects described in the main text, with attention and Predictability being the only factors showing significant effects (see Table 19).

Table 19. Model parameters for the student sample.

	X^2	*df*	*p*
(Intercept)	939.614	1.000	<0.001*
Attention	21.350	1.000	<0.001*
Predictability	21.644	2.000	<0.001*
AQ	3.568	1.000	0.059
Attention X Predictability	4.238	2.000	0.120
Attention x AQ	0.216	1.000	0.642
Predictability x AQ	1.328	2.000	0.515
Attention x Predictability x AQ	0.975	2.000	0.614

***Note*: * indicate statistically significant effects.**

Like in the overall analysis, Attended targets were detected faster than Unattended targets – β = 0.059[SE=0.013, 0.034 - 0.083], z-ratio = 4.621, p < 0.001. Additionally, Expected targets were detected faster than Unexpected targets - β = 0.089[SE=0.019, 0.043 – 0.135], z-ratio = 4.652, p < 0.001. When no expectation was set, targets were detected faster than when the targets were

unexpected - β = 0.072[SE=0.020, 0.024 – 0.121], z-ratio = 3.544, p = 0.001, however Expected targets were not detected faster than when no expectation was set - β = 0.017[SE=0.014, -0.016 – 0.049], z-ratio = 1.216, p = 0.532. These results are consistent with the effects observed in the combined analysis.

Prolific sample

In Table 20, we can see that like the full and the student sample, in the Prolific sample variations between different conditions are small for accuracy indicating that individuals were able to perform the task. Additionally, reaction time appears faster for Expected targets in comparison to Unexpected targets. Further, looking at the Expected and the Unexpected conditions, Attended targets elicited shorter reaction times in comparison to unattended target. The opposite was observed in the None condition, where Attended targets elicited longer reaction times, which is an indication of potential inhibition of return. This difference with attention is not observed at the student sample.

Table 20 Participant accuracy and median reaction time

Attention	Predictability	Mean accuracy (SD)	Median RT (MAD)*
Attended	Expected	0.943 (0.232)	0.497 (0.096)
	None	0.943 (0.232)	0.507 (0.111)
	Unexpected	0.936 (0.245)	0.516 (0.102)
Unattended	Expected	0.935 (0.247)	0.502 (0.105)
	None	0.937 (0.243)	0.500 (0.107)
	Unexpected	0.925 (0.264)	0.528 (0.117)

***Note*: * MAD = Median absolute deviation.**

The overall results from the models are consistent with the overall effects described in the main text, with Attention and Predictability being the only factors showing significant effects (see Table 21). Attended targets were detected faster than Unattended targets, however the CIs were broader than in the other two samples potentially driven by the observed IOR effects in the None condition – β = 0.023[SE=0.012, 0.0006 - 0.046], z-ratio = 2.020, p = 0.043. Looking at the simple effects of Predictability, the effects mirrored the overall and the student samples. There was no significant difference in the speed of detection between expected and None targets- β = 0.020[SE=0.012, -0.010 – 0.049], z-ratio = 1.611, p = 0.288. None targets, however, were detected faster than Unexpected targets - β = 0.065[SE=0.013, 0.033 – 0.097], z-ratio = 4.876, p

< 0.001, and Expected targets were detected faster than Unexpected targets – *B* = 0.085[SE=0.018, 0.043 – 0.127], z-ratio = 4.809, *p* < 0.001.

Table 21. Model estimates for the Prolific sample.

	X^2	*df*	*p*
(Intercept)	879.712	1.000	<0.001*
Attention	4.078	1.000	0.043*
Predictability	27.206	2.000	<0.001*
AQ	0.644	1.000	0.422
Attention X Predictability	2.750	2.000	0.253
Attention x AQ	0.535	1.000	0.464
Predictability x AQ	0.293	2.000	0.864
Attention x Predictability x AQ	0.439	2.000	0.803

***Note*: * indicate statistically significant effects.**

It is noteworthy that while the models for the students was equivalent to the model used for the full sample, due to further singularity problems, the gamma model for the Prolific sample, contained only intercept for participants and intercept for grating direction, however, it maintained the full random effect structure for the participant X grating direction.

Coherent motion

Student sample

In Table 22 is reported the data for the student and Prolific samples. Variations between different conditions are small for accuracy indicating that individuals were able to do the task. Additionally, reaction time appears faster for Expected targets in comparison to Unexpected targets. Further, looking at the Unexpected and the None conditions, Attended targets elicited longer reaction times in comparison to Unattended target, which is an indication of inhibition of return, which was not observed in the BM experiment. This difference in attention is not present for the Expected condition.

Table 22 Participant accuracy and median reaction time

Attention	Predictability	Mean accuracy (SD)	Median RT (MAD)*
Attended	Expected	0.913 (0.282)	0.502 (0.109)
	None	0.900 (0.300)	0.515 (0.121)
	Unexpected	0.871 (0.336)	0.541 (0.114)
Unattended	Expected	0.909 (0.288)	0.502 (0.119)
	None	0.896 (0.305)	0.502 (0.115)
	Unexpected	0.906 (0.292)	0.533 (0.125)

***Note*: * MAD = Median absolute deviation.**

Similar to the combined sample and the Prolific BM experiment samples that exhibited IOR in the None condition, there was no significant effect of Attention, but the effect of Predictability was retained (see Table 23). Similar to the combined CM results and the BM experiment, Expected targets were not detected faster than None targets - β = 0.020[SE=0.014, -0.013 – 0.059], z-ratio = 1.430, p = 0.392. However, targets where no expectation was set were detected faster than Unexpected targets - β = 0.097[SE=0.019, 0.052 – 0.143], z-ratio = 5.107, p < 0.001, and Expected targets were detected faster than Unexpected targets - β = 0.117[SE=0.020, 0.068 – 0.166], z-ratio = 5.724, p < 0.001.

Table 23. Model estimated for the student sample.

	X^2	*df*	*p*
(Intercept)	1002.670	1.000	<0.001*
Attention	1.202	1.000	0.273
Predictability	34.193	2.000	<0.001*
AQ	0.051	1.000	0.822
Attention X Predictability	0.998	2.000	0.607
Attention x AQ	2.213	1.000	0.137
Predictability x AQ	0.013	2.000	0.994
Attention x Predictability x AQ	2.768	2.000	0.251

***Note*: * indicate statistically significant effects.**

Prolific sample

Observing the data in Table 24, variations between different conditions are small for accuracy indicating that individuals were able to do the task. Additionally, reaction time appears faster for Expected targets in comparison to Unexpected targets. Further, looking at the Unexpected and the None conditions, Attended targets elicited longer reaction times in comparison to Unattended target, which is an indication of inhibition of return, which was not observed in the BM experiment. This difference with attention is also not observed at the Expected condition.

Table 24. Participant accuracy and median reaction time

Attention	Predictability	Mean accuracy (SD)	Median RT (MAD)*
Attended	Expected	0.937 (0.243)	0.498 (0.100)
	None	0.919 (0.273)	0.502 (0.106)
	Unexpected	0.912 (0.284)	0.523 (0.112)
Unattended	Expected	0.929 (0.256)	0.499 (0.101)
	None	0.922 (0.268)	0.495 (0.104)
	Unexpected	0.931 (0.254)	0.520 (0.106)

***Note*: * MAD = Median absolute deviation.**

Like the student sample and the combined sample, the Prolific sample showed no significant main effect of attention, but a significant effect of Predictability and no other significant effects or interactions (see Table 25). There was no significant difference in the speed of detection between Expected and None targets- β = 0.002[SE=0.014, -0.032 – 0.036], z-ratio = 0.161, p = 0.998. The targets in the None condition were detected faster than Unexpected targets – β = 0.098[SE=0.019, 0.053 – 0.143], z-ratio = 5.239, p < 0.001, and Expected targets were detected faster than Unexpected targets - β = 0.101[SE=0.019, 0.057 – 0.145], z-ratio = 5.412, p < 0.001.

Table 25. Model estimates for the Prolific sample.

	X^2	*df*	*p*
(Intercept)	1357.793	1.000	<0.001*
Attention	0.099	1.000	0.753
Predictability	33.269	2.000	<0.001*
AQ	0.010	1.000	0.922
Attention X Predictability	3.543	2.000	0.170
Attention x AQ	0.659	1.000	0.417
Predictability x AQ	2.478	2.000	0.290
Attention x Predictability x AQ	0.502	2.000	0.778

***Note*: * indicate statistically significant effects.**

It is noteworthy that unlike the other Prolific samples, the gamma model here differed from the model in the main text by only removing the slopes for the participant random factor with the rest of the random factor structure identical.

Biological motion findings replication

Due to researcher error, a Prolific sample with a limited age of 18 was also recruited. As the sample was not representative of the age distribution, the analysis of this Prolific sample is only used to show the replicability of the findings. We also combined all the participants to further increase power and emphasise the stability of the results.

Combined samples analysis

After cleaning and combining the 3 samples (1 student and 2 Prolific), we were left with 101 participants (Mage = 19.366(SD =2.284), 46 male/ 55 female, AQ = 3.762 (1.664). In Table 26 are reported the RT and accuracy data for each of the factor levels of Attention and Predictability for the whole sample. Reaction time

appears faster for Expected targets in comparison to Unexpected targets. Further, Attended targets elicited shorter reaction times in comparison to Unattended target. However, there appears to be no difference in the None condition between attended and Unattended targets, suggesting that the attention cuing of the PLD was not sufficient to influence behaviour. These effects are more easily observed Figure 27 and Figure 28. However, we can see that the observed trends in the data with increase in AQ from the main analysis remain.

Table 26 Participant accuracy and median reaction time

Attention	Predictability	Mean accuracy (SD)	Median RT (MAD)*
Attended	Expected	0.925 (0.263)	0.503 (0.106)
	None	0.913 (0.282)	0.518 (0.121)
	Unexpected	0.893 (0.309)	0.533 (0.122)
Unattended	Expected	0.914 (0.281)	0.518 (0.121)
	None	0.908 (0.289)	0.518 (0.121)
	Unexpected	0.897 (0.304)	0.543 (0.135)

***Note*: * MAD = Median absolute deviation.**

Figure 27. Reaction time for accurately detected targets by attention and predictability conditions.

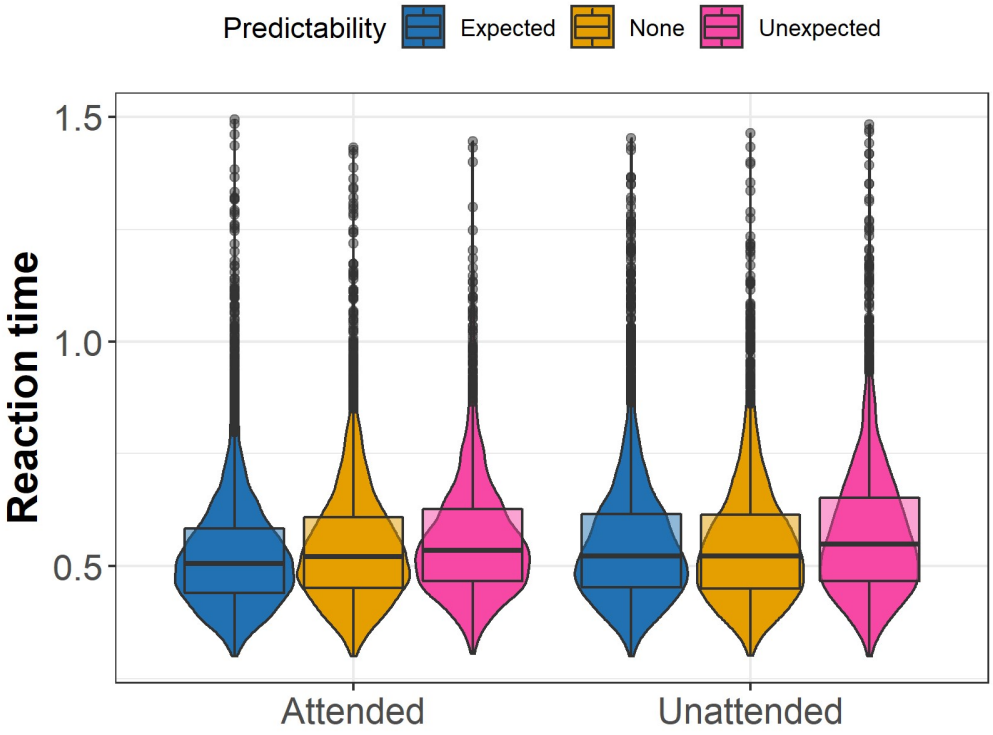
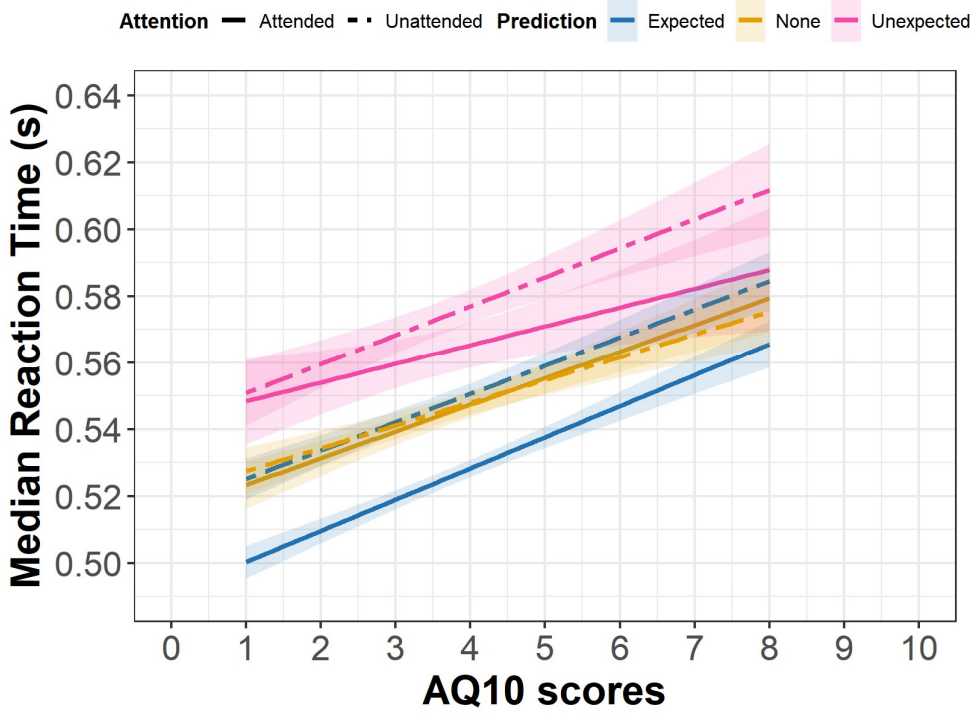


Figure 28. Changes in RT with AQ scores for Attention (line types) and Prediction levels (colours).



Overall, there was a significant main effect of Attention ($X^2(1) = 28.872$, $p < 0.001$) and Predictability ($X^2(2) = 41.929$, $p < 0.001$). However, there was not a significant interaction between the two variables ($X^2(2) = 5.160$, $p = 0.076$). Unlike the main analysis, the addition of more participants resulted in the effect of AQ becoming significant($X^2(1) = 4.084$, $p = 0.043$), where participants with higher AQ had overall longer RTs. However, there were again no significant 2-way interaction with Predictability ($X^2(2) = 0.931$, $p = 0.628$)), no significant 2-way interaction with Attention ($X^2(1) = 0.002$, $p = 0.961$) and no significant 3-way interactions ($X^2(2) = 1.305$, $p = 0.521$). The interaction between Attention and Platform was also still significant ($X^2(1) = 5.043$, $p = 0.025$). No other effects or interactions were significant with Platform – all p-values > 0.138, indicating that there were no other differences between the different platforms.

Overall, Attended targets were detected faster than Unattended ones - $B = 0.041$ [SE=0.008, 0.026 – 0.056], z-ratio = 5.373, $p < 0.001$. Additionally, as seen in Table 27, students showed a bigger difference in response time between attended and Unattended targets, in comparison to the participants recruited through Prolific. Although, unlike in the main analysis, the addition of the

second Prolific sample has increased the effect between Attended and Unattended targets making the difference significant for both platforms.

Table 27. Simple effects of attention at each platform level

Contrast	Platform	estimate	SE	LCI	UCI	*Z*-ratio	*p*-value
attended - unattended	Prolific	0.024	0.010	0.003	0.046	2.531	0.023*
attended - unattended	Student	0.058	0.012	0.032	0.084	4.953	<0.001*

***Note*: * indicate statistically significant contrast.**

Adding the additional Prolific participants did not change the simple contrasts between Predictability levels observed in the main analysis. Firstly, Expected targets were detected faster than None targets, however this difference was still not significant despite the larger sample - β = 0.018[SE=0.009, -0.003 – 0.041], z-ratio = 2.034, p = 0.121. None targets, were still detected faster than Unexpected targets - β = 0.070[SE=0.013, 0.040 – 0.101], z-ratio = 5.527, p < 0.001, and Expected targets were detected faster than Unexpected targets - β = 0.089[SE=0.014, 0.056 – 0.122], z-ratio = 6.397, p < 0.001.

Thus, as expected, predictability assists in participants ability to respond faster to Expected targets, however, that ability is not modulated by attention or AQ scores. However, the addition of a few more participants pushed the main effect of AQ over the chosen significance threshold, indicating that the effect AQ is weak, but present. This suggests that the effect might be more pronounced in autistic individuals. This is further in line with the increased processing time for autistic individuals reported in previous research (eg. Lawson et al., 2014).

Prolific sample only

After cleaning, we were left with 31 participants (M_{age} = 18.032(0.180), 23 male 8 female, AQ = 3.868 (1.408), MacOS =1, Windows = 30, Chrome = 22, Mozilla Firefox = 3, other = 6). Observing the data in Table 28Table 24, variations between different conditions are small for accuracy indicating that individuals were able to perform the task. Additionally, reaction time appears faster for Expected targets in comparison to Unexpected targets. Further, looking at the Expected and the Unexpected conditions, Attended targets elicited shorter reaction times in comparison to Unattended target. The opposite was observed

in the None condition, where Attended targets elicited longer reaction times, which is an indication of potential inhibition of return which was observed in the CM experiment in both samples and in the other Prolific sample in the BM experiment. This difference with attention is not observed at the student or the combined samples for the BM experiment.

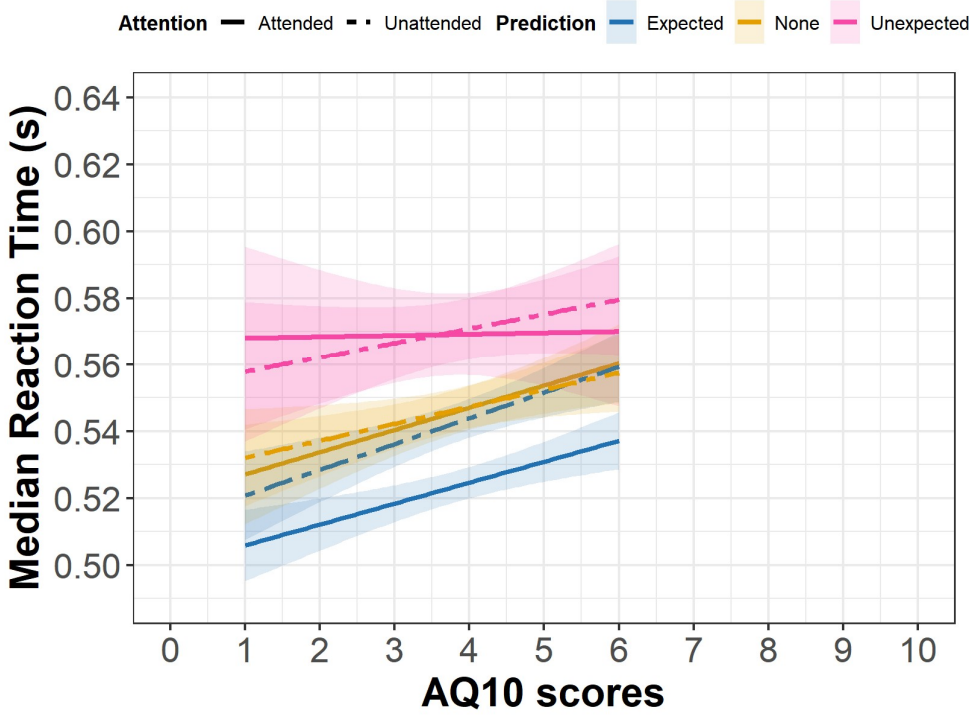
Table 28 Participant accuracy and median reaction time

Attention	Predictability	Mean accuracy (SD)	Median RT (MAD)*
Attended	Expected	0.919 (0.272)	0.503 (0.106)
	None	0.904 (0.294)	0.520 (0.117)
	Unexpected	0.892 (0.311)	0.538 (0.127)
Unattended	Expected	0.898 (0.302)	0.516 (0.119)
	None	0.903 (0.296)	0.519 (0.118)
	Unexpected	0.882 (0.323)	0.541 (0.135)

***Note*: * MAD = Median absolute deviation.**

Looking at the effects of AQ on performance, we observe similar trends as in the main analysis (see Figure 29). There appears to be an increase in RT with increase in AQ, although that is not the case for attended Unexpected targets. Additionally, we can see that we observe an IOR for Unexpected targets at lower AQ scored, but not at higher AQ scores. Moreover, in the None condition we only see an IOR at higher AQ levels, which was evident in the main sample. However, since this was also seen in the main CM experiment, it could be that higher AQ scores are only diminishing the difference between attended and Unattended targets.

Figure 29. Changes in RT with AQ scores for Attention (line types) and Prediction levels (colours).



Overall, unlike the other Prolific sample, there was an effect of Attention. Additionally, the Predictability effect is still evident (see Table 29).

Table 29. Model estimates for the second Prolific sample.

	X^2	*df*	*p*
(Intercept)	838.657	1.000	<0.001*
Attention	5.474	1.000	0.019*
Predictability	17.050	2.000	<0.001*
AQ	0.335	1.000	0.563
Attention X Predictability	0.308	2.000	0.857
Attention x AQ	0.000	1.000	0.988
Predictability x AQ	0.738	2.000	0.692
Attention x Predictability x AQ	2.854	2.000	0.240

***Note*: * indicate statistically significant effects.**

Despite the observed IOR in the None condition in Table 28, Attended targets were detected faster than Unattended ones - β = 0.030 [SE=0.013, 0.005 – 0.056], z-ratio = 2.340, p = 0.0193. Like in all other samples, there was no significant difference in the speed of detection between Expected and None targets- β = 0.028[SE=0.018, -0.014 – 0.070], z-ratio = 1.605, p = 0.291. None targets were, however, detected faster than Unexpected targets - β = 0.072[SE=0.023, 0.018 – 0.127], z-ratio = 3.176, p = 0.005, and Expected targets

were detected faster than Unexpected targets - β = 0.100[SE=0.024, 0.042 – 0.159], z-ratio = 4.120, p < 0.001.

It is noteworthy that similar to the other Prolific sample in the main experiment, there were further singularity problems. Thus, the gamma model contained only intercept for participants and intercept for grating direction, however, it maintained the full random effect structure for the participant X grating direction.

Combined analysis for the BM and CM experiment

To be able to say with some degree of certainty, that there was a difference in the BM and the CM experiments which was caused by the type of stimulus rather than the social component of the stimulus, we ran a model combining the experiments reported in the main text, including experiment as an additional fixed factor. As seen in Table 30 and in Table 31 there was a significant interaction between Attention and Experiment. This was due to larger differences between the Attended and Unattended trials in the BM experiment, and it also indicates a non-significant but reversed response times for Attended and Unattended targets in the CM experiment. This can also be seen in Figure 30. importantly, there is no interaction with AQ indicating that any effects of the attention modulation between the two samples was not due to differences in the level of autistic traits. Thus, we do not believe that the differences in results between the two experiments is caused by the social/non-social nature of the tasks.

Table 30. Model estimates for the combined experiments reported in the main text

	X^2	*df*	*p*
(Intercept)	2714.070	1.000	0.000
Attention	5.891	1.000	0.015*
Expectation	86.115	2.000	0.000*
AQ	2.214	1.000	0.137
Platform	2.550	1.000	0.110
Experiment	0.169	1.000	0.681
Attention X Expectation	5.119	2.000	0.077
Attention x AQ	1.954	1.000	0.162
Expectation x AQ	0.304	2.000	0.859
Attention X Platform	1.834	1.000	0.176
Expectation x Platform	0.552	2.000	0.759
AQ x Platform	0.231	1.000	0.631
Attention X Experiment	14.769	1.000	<0.001*
Expectation x Experiment	2.420	2.000	0.298
Experiment x AQ	1.312	1.000	0.252
Experiment x Platform	0.302	1.000	0.583
Attention x Expectation x AQ	2.745	2.000	0.253
Attention x Expectation x Platform	0.939	2.000	0.625
Attention x AQ x Platform	0.005	1.000	0.946
Expectation x AQ x Platform	0.931	2.000	0.628
Attention x Expectation x Experiment	1.376	2.000	0.503
Attention x AQ x Experiment	1.248	1.000	0.264
Expectation x AQ x Experiment	0.823	2.000	0.663
Attention x Platform x Experiment	2.615	1.000	0.106
Expectation x Platform x Experiment	0.183	2.000	0.913
AQ x Platform x Experiment	0.132	1.000	0.717
Attention x Expectation x AQ x Platform	1.244	2.000	0.537
Attention x Expectation x AQ x Experiment	0.211	2.000	0.900
Attention x Expectation x Platform x Experiment	1.857	2.000	0.395
Attention x AQ x Platform x Experiment	1.155	1.000	0.283
Expectation x AQ x Platform x Experiment	1.111	2.000	0.574
Attention x Expectation x AQ x Platform x Experiment	1.598	2.000	0.450

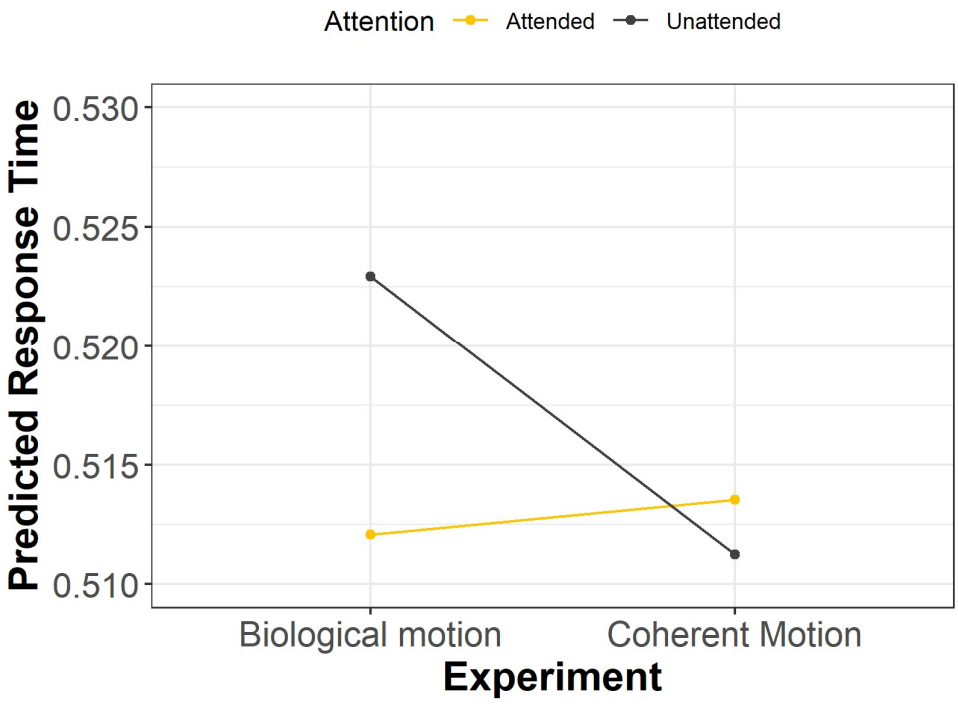
***Note*: * indicate statistically significant effects.**

Table 31. Simple effects of Attention for each Experiment

Contrast	Experiment	estimate	SE	LCI	UCI	*Z*-ratio	*p*-value
attended-unattended	BM	0.041	0.009	0.020	0.061	4.369	<0.001*
attended-unattended	CM	-0.009	0.009	-0.029	0.011	-0.970	0.554

***Note*: * indicate statistically significant effects.**

Figure 30. Differences between Attention levels for the BM and CM experiments.



Appendix E

Pilot studies' design and results for Chapter 5

To make sure that the participants were able to perform the task described in Chapter 5 and to select the appropriate amount of noise for the main experiment, two pilot studies were designed. To maintain similar difficulty conditions to the main experiment, the PLD and surrounding noise flickered at 5Hz frequency, equivalent to the SSVEP frequency chosen below.

The first pilot varied the number of noise dots surrounding a point-light display of a walking human (the 'dots' pilot). The aim was to allow us to select the appropriate amount of noise. The second pilot tested how many frames are necessary for an individual to be able to detect the direction of walking of the point-light display (the 'frames' pilot). The 'dots' pilot was run first; the results were analysed and then the 'frames' pilot was run using the parameters from the 'dots' pilot.

'Dots' pilot

Eight participants took part in the 'dots' pilot. One participant did not follow the instructions/did not pay attention, and for another, the paradigm malfunctioned. The remaining 6 participants had a mean age of 28.33 (SD = 3.88) and there was an equal split between males and females.

There were 10 different levels of noise. All participants started with no-noise – i.e., only the scrambled walker that was representing the direction that was not coherent at the given point. The no-noise condition appeared at random throughout the paradigm to serve as baseline observation. The paradigm was set up for there to be 10 trials per condition, however, due to technical difficulties (introduction of lag in presentation by the end of data collection due to lack of RAM space), there was a varying amount of data per condition[5]. Stimuli were presented on a CRT screen at 120Hz at the centre of the screen at a visual distance of ~ 75cm.

[5] This was also the case for the 'Frames' pilot reported after.

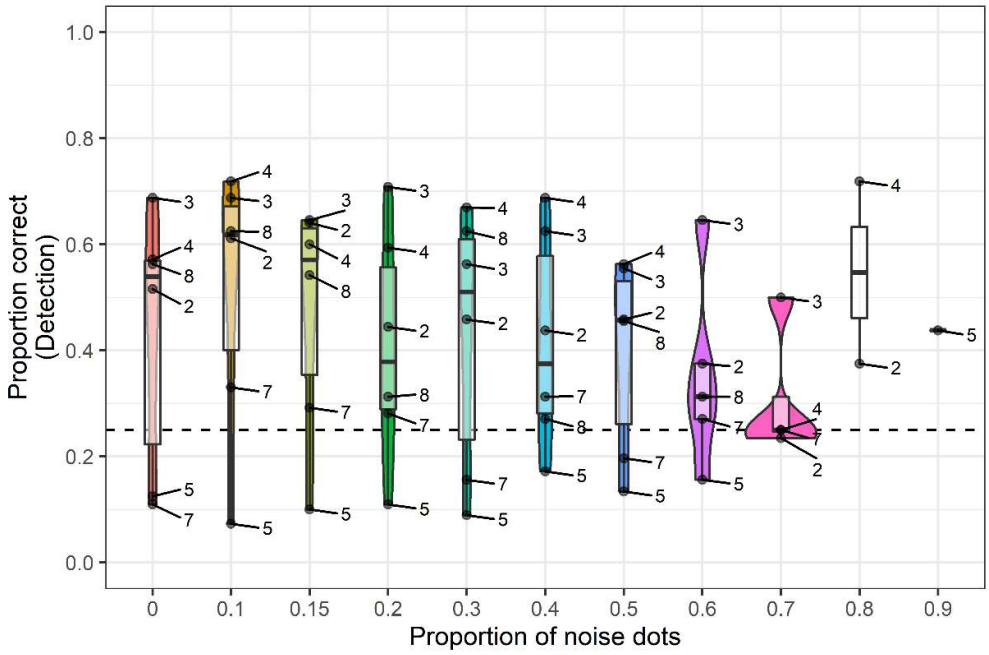
Participants were asked to indicate the direction of walking on the PLD every time they detected one in a trial. Accuracy was solely measured based on detection and not on the direction of walking at this point. There were 16 coherent walkers in each trial, and each was fully coherent for 31 frames. Each participant's accuracy was calculated based on the total number of trials and conditions they saw. Accuracy was calculated for button presses which 1) occurred 200ms after the last frame of the coherent walker, 2) were the first response, and 3) were not considered as an outlier in each individual's reaction time distribution, by using the inter-quartile method.

All participants showed low detection rates regardless of the amount of noise. It is also noteworthy that chance performance here would not be 50%, as the participants needed firstly to be able to detect the PLD which has a 50% probability (seen/not seen), and then they had to correctly identify the direction of walking. As these probabilities are independent the chance detection rate would be 0.5*0.5 = 0.25, i.e., 25%. The best performance was observed when the noise was created by using only 1-2 dots from the six scrambled walkers (see Table 32Error! Reference source not found. & Figure 31). The 80% and 90% detection rates were not taken into account because they represented very few participants due to the data loss. The most commonly occurring dots in the correct trials, that also maintained some movement from frame to frame, were the light dots representing the left wrist and the right ankle joint. These were used for the 'frames' pilot described below and the final study.

Table 32. Proportion detected per condition.

Proportion of noise dots	Number of noise dots (x6)	Proportion detected (SD)
0.00	0	0.43 (0.25)
0.10	1	0.51 (0.25)
0.15	2	0.47(0.22)
0.20	3	0.41 (0.22)
0.30	4	0.43 (0.25)
0.40	5	0.42 (0.20)
0.50	6 or 7	0.39 (0.18)
0.60	8	0.35 (0.18)
0.70	9	0.31 (0.13)
0.80	10 or 11	0.55 (0.24)
0.90	12	0.44 (NA)

Figure 31. Proportion of correctly detected PLDs per participant and condition. (Dashed line represents 25%). Dots represent individuals' proportion correct.



It is noteworthy that participants tended to perform worse when there was 'no noise' (i.e., only the additional scrambled walker was present). We presume that this is due to the number of dots being in the same place as the coherent walker, making it more difficult to distinguish the coherent walker from the noise. In fact, a study by Trevino et al. (2016) showed a U-shaped function between the amount of noise and accuracy. Specifically, the inclusion of moderate noise in their study improved visual motion detection. We believe that this same mechanism might have come into play with our pilot.

'Frames' pilot

In this pilot we varied the length of time (the number of frames) that the coherent walker was presented for – i.e., we varied the plateau of the sine curve of the PLD stimulus. In the 'dots' pilot, the coherent walker was present for 31 frames at 120Hz, which was equivalent to ~258ms. As participants' performance was rarely above 70% in that pilot, here we used 8 conditions, in which the length of the coherent walker presentation varied from 17 to 45 frames in steps of 4.

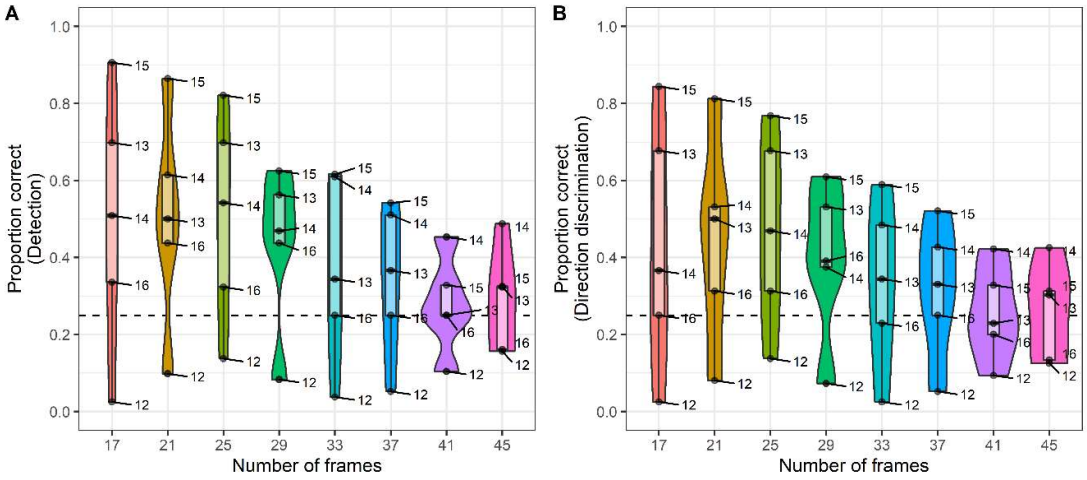
Participants were again asked to indicate the direction of the walking PLD. Five participants took part in this pilot 1 female and 4 males, with a mean age of

24.4 (SD = 3.97). One participant took part in both pilots. We analysed accuracy in detecting the direction of the PLD. Surprisingly, the more frames we provided (the longer the coherent PLD presentation) led to lower accuracy. We also looked at detection rates to verify that participants were not just guessing but differences were very minimal on the sample level (see Table 33). The participants that showed high detection rates, were also performing better in detecting the direction of walking (see Figure 32 below). There was wide between-participant variability, thus, we initially decided to keep the 31 frames at 120Hz refresh rate. Due to equipment availability and script functionality, in the final study, we ended up using 17 frames at a 60Hz refresh rate, which is equivalent to ~284ms. It is noteworthy that this is ~26ms longer than the decided 31 frames at 120Hz. Based on the results that we obtained in the two pilot studies and the changes to the paradigm that we needed to make, we incorporated a training procedure to bring the participants to an accuracy of at least 65% for the experimental condition.

Table 33. Proportion detected and correctly identified the direction of walking per condition.

Number of frames	Proportion detected (SD)	Proportion detected direction (SD)
17	0.49 (0.34)	0.43 (0.33)
21	0.50 (0.28)	0.45 (0.27)
25	0.50 (0.28)	0.47 (0.26)
29	0.44 (0.21)	0.40 (0.21)
33	0.37 (0.25)	0.33 (0.22)
37	0.34 (0.20)	0.32 (0.18)
41	0.28 (0.13)	0.25 (0.13)
45	0.29 (0.14)	0.26 (0.13)

Figure 32. Responses of participants from the 'frames' pilot.



***Note*: A - Detection rates of participants. B - Accuracy in walking direction discrimination. (Dashed line represents 25%). Dots represent each individual's proportion correct.**

References

Actis-Grosso, R., Bossi, F., & Ricciardelli, P. (2015). Emotion recognition through static faces and moving bodies: a comparison between typically developed adults and individuals with high level of autistic traits. *Frontiers in Psychology*, *6*, 1570. https://doi.org/10.3389/fpsyg.2015.01570

Alaerts, K., Geerlings, F., Herremans, L., Swinnen, S. P., Verhoeven, J., Sunaert, S., & Wenderoth, N. (2015). Functional Organization of the Action Observation Network in Autism: A Graph Theory Approach. *PLoS ONE*, *10*(8), 1–21. https://doi.org/10.1371/journal.pone.0137020

Alaerts, K., Swinnen, S. P., & Wenderoth, N. (2016). Sex differences in autism: a resting-state fMRI investigation of functional brain connectivity in males and females. *Social Cognitive and Affective Neuroscience*, *11*(6), 1002–1016. https://doi.org/10.1093/scan/nsw027

Alaerts, K., Swinnen, S. P., & Wenderoth, N. (2017). Neural processing of biological motion in autism: An investigation of brain activity and effective connectivity. *Scientific Reports*, *7*(1), 5612. https://doi.org/https://dx.doi.org/10.1038/s41598-017-05786-z

Alaerts, K., Woolley, D. G., Steyaert, J., Di Martino, A., Swinnen, S. P., & Wenderoth, N. (2013). Underconnectivity of the superior temporal sulcus predicts emotion recognition deficits in autism. *Social Cognitive and Affective Neuroscience*, *9*(10), 1589–1600. https://doi.org/10.1093/scan/nst156

Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus Predictability Reduces Responses in Primary Visual Cortex. *Journal of Neuroscience*, *30*(8), 2960–2966. https://doi.org/10.1523/JNEUROSCI.3730-10.2010

Allenmark, F., Shi, Z., Pistorius, R. L., Theisinger, L. A., Koutsouleris, N., Falkai, P., Müller, H. J., & Falter-Wagner, C. M. (2020). Acquisition and Use of 'Priors' in Autism: Typical in Deciding Where to Look, Atypical in Deciding

What Is There. *Journal of Autism and Developmental Disorders*, 1–15. https://doi.org/10.1007/s10803-020-04828-2

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward Brief "Red Flags" for Autism Screening: The Short Autism Spectrum Quotient and the Short Quantitative Checklist in 1,000 Cases and 3,000 Controls. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(2), 202-212.e7. https://doi.org/10.1016/j.jaac.2011.11.003

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/https://doi.org/10.1176/appi.books.9780890425596

Ames, C., & Fletcher-Watson, S. (2010). A review of methods in the study of attention in autism. *Developmental Review*, *30*(1), 52–73. https://doi.org/10.1016/j.dr.2009.12.003

Amoruso, L., Narzisi, A., Pinzino, M., Finisguerra, A., Billeci, L., Calderoni, S., Fabbro, F., Muratori, F., Volzone, A., & Urgesi, C. (2019). Contextual priors do not modulate action prediction in children with autism. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1908), 20191319. https://doi.org/10.1098/rspb.2019.1319

Annaz, D., Campbell, R., Coleman, M., Milne, E., & Swettenham, J. (2012). Young children with autism spectrum disorder do not preferentially attend to biological motion. *Journal of Autism and Developmental Disorders*, *42*(3), 401–408. https://doi.org/10.1007/s10803-011-1256-3

Annaz, D., Remington, A., Milne, E., Coleman, M., Campbell, R., Thomas, M. S. C., & Swettenham, J. (2010). Development of motion processing in children with autism. *Developmental Science*, *13*(6), 826–838. https://doi.org/10.1111/j.1467-7687.2009.00939.x

Antezana, L., Factor, R. S., Condy, E. E., Strege, M. V., Scarpa, A., & Richey, J. A. (2019). Gender differences in restricted and repetitive behaviors and interests in youth with autism. *Autism Research*, *12*(2), 274–283.

https://doi.org/10.1002/aur.2049

Anwyl-Irvine, A., Dalmaijer, E. S., Hodges, N., & Evershed, J. K. (2020). Online Timing Accuracy and Precision: A comparison of platforms, browsers, and participant's devices. *PsyArXiv*. https://pdfs.semanticscholar.org/3ccf/fb46a2c05131a6086ad7bf86edb7e8f904ef.pdf?_ga=2.252001611.31985767.1590567086-1250328828.1590410032

ASDEU. (2018). *Autism Spectrum Disorders in the European Union: Executive Summary*. http://asdeu.eu/wp-content/uploads/2016/12/ASDEUExecSummary27September2018.pdf

Atkinson, A. P. (2009). Impaired recognition of emotions from body movements is associated with elevated motion coherence thresholds in autism spectrum disorders. *Neuropsychologia*, *47*(13), 3023–3029. https://doi.org/10.1016/j.neuropsychologia.2009.05.019

Atkinson, M. A., Simpson, A. A., & Cole, G. G. (2018). Visual attention and action: How cueing, direct mapping, and social interactions drive orienting. *Psychonomic Bulletin & Review*, *25*(5), 1585–1605. https://doi.org/10.3758/s13423-017-1354-0

Bach, M. (2007). The Freiburg Visual Acuity Test-Variability unchanged by post-hoc re-analysis. *Graefe's Archive for Clinical and Experimental Ophthalmology*, *245*(7), 965–971. https://doi.org/10.1007/s00417-006-0474-4

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., Kurzius-Spencer, M., Zahorodny, W., Robinson Rosenberg, C., White, T., Durkin, M. S., Imm, P., Nikolaou, L., Yeargin-Allsopp, M., Lee, L.-C., Harrington, R., Lopez, M., Fitzgerald, R. T., Hewitt, A., … Dowling, N. F. (2018). Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014. *Morbidity and Mortality Weekly Report. Surveillance Summaries (Washington, D.C. : 2002)*, *67*(6), 1–23. https://doi.org/10.15585/mmwr.ss6706a1

Bal, E., Harden, E., Lamb, D., Van Hecke, A. V., Denver, J. W., & Porges, S. W. (2010). Emotion Recognition in Children with Autism Spectrum Disorders: Relations to Eye Gaze and Autonomic State. *Journal of Autism and Developmental Disorders*, *40*(3), 358–370. https://doi.org/10.1007/s10803-009-0884-3

Baron-Cohen, S. (2000). Theory of mind and autism: A review. *International Review of Research in Mental Retardation*, *23*, 169–184. https://doi.org/10.1016/S0074-7750(00)80010-5

Baron-Cohen, S., Allen, J., & Gillberg, C. (1992). Can Autism be Detected at 18 Months? *British Journal of Psychiatry*, *161*(6), 839–843. https://doi.org/10.1192/bjp.161.6.839

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

Barrett, S. L., Uljarević, M., Jones, C. R. G., & Leekam, S. R. (2018). Assessing subtypes of restricted and repetitive behaviour using the Adult Repetitive Behaviour Questionnaire-2 in autistic adults. *Molecular Autism*, *9*(1), 1–10. https://doi.org/10.1186/s13229-018-0242-4

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). *Parsimonious Mixed Models*. http://arxiv.org/abs/1506.04967

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

Bernier, R., Aaronson, B., & Mcpartland, J. (2013). The role of imitation in the observed heterogeneity in EEG mu rhythm in autism and typical development. *Brain and Cognition*, *82*, 69–75. https://doi.org/10.1016/j.bandc.2013.02.008

Bernier, R., Dawson, G., Webb, S., & Murias, M. (2007). EEG mu rhythm and imitation impairments in individuals with autism spectrum disorder. *Brain and Cognition*, *64*(3), 228–237. https://doi.org/10.1016/j.bandc.2007.03.004

Bertone, A., Mottron, L., Jelenic, P., & Faubert, J. (2003). Motion Perception in Autism: A "Complex" Issue. *Journal of Cognitive Neuroscience*, *15*(2), 218–225. https://doi.org/10.1162/089892903321208150

Binnersley, J. (2006). *Perception of Biological Motion in Autistic Spectrum Disorder* [University College London]. http://discovery.ucl.ac.uk/1569672/1/Binnersley_thesis.pdf

Bird, G., & Cook, R. (2013). Mixed emotions: the contribution of alexithymia to the emotional symptoms of autism. *Translational Psychiatry 2013 3:7*, *3*(7), e285–e285. https://doi.org/10.1038/tp.2013.61

Björnsdotter, M., Wang, N., Pelphrey, K., & Kaiser, M. D. (2016). Evaluation of quantified social perception circuit activity as a neurobiological marker of autism spectrum disorder. *JAMA Psychiatry*, *73*(6), 614–621. https://doi.org/10.1001/jamapsychiatry.2016.0219

Blake, R., & Shiffrar, M. (2007). Perception of Human Motion. *Annual Review of Psychology*, *58*(1), 47–73. https://doi.org/10.1146/annurev.psych.57.102904.190152

Blake, R., Turner, L. M. M., Smoski, M. J. J., Pozdol, S. L. L., & Stone, W. . L. (2003). Visual recognition of biological motion is impaired in children with autism. *Psychol Sci*, *14*(2), 151–157. https://doi.org/10.1111/1467-9280.01434

Boxhoorn, S., Bast, N., Supèr, H., Polzer, L., Cholemkery, H., & Freitag, C. M. (2020). Pupil dilation during visuospatial orienting differentiates between autism spectrum disorder and attention-deficit/hyperactivity disorder. *Journal of Child Psychology and Psychiatry*, *61*(5), 614–624. https://doi.org/10.1111/jcpp.13179

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Bridges, D., Pitiot, A., MacAskill, M., & Peirce, J. (2020). The timing mega-study: comparing a range of experiment generators, both lab-based and online. *PsyArXiv*. https://psyarxiv.com/d6nu5/

Brock, J. (2012). Alternative Bayesian accounts of autistic perception: Comment on Pellicano and Burr. In *Trends in Cognitive Sciences* (Vol. 16, Issue 12, pp. 573–574). Elsevier. https://doi.org/10.1016/j.tics.2012.10.005

Brosnan, M. J., Gwilliam, L. R., & Walker, I. (2012). Brief report: The relationship between visual acuity, the embedded figures test and systemizing in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *42*(11), 2491–2497. https://doi.org/10.1007/s10803-012-1505-0

Brugha, T., Cooper, S., McManus, S., Purdon, S., Smith, J., Scott, F., Spiers, N., & Tyrer, F. (2012). *Estimating the Prevalence of Autism Spectrum Conditions in Adults: Extending the 2007 Adult Psychiatric Morbidity Survey*. https://files.digital.nhs.uk/publicationimport/pub05xxx/pub05061/esti-prev-auti-ext-07-psyc-morb-surv-rep.pdf

Brugha, T., Tyrer, F., Leaver, A., Lewis, S., Seaton, S., Morgan, Z., Tromans, S., & van Rensburg, K. (2020). Testing adults by questionnaire for social and communication disorders, including autism spectrum disorders, in an adult mental health service population. *International Journal of Methods in Psychiatric Research*, *29*(1), e1814. https://doi.org/10.1002/mpr.1814

Brysbaert, M. (2019). How Many Participants Do We Have to Include in Properly Powered Experiments? A Tutorial of Power Analysis with Reference Tables. *Journal of Cognition*, *2*(1). https://doi.org/10.5334/joc.72

Burnside, K., Wright, K., & Poulin-Dubois, D. (2017). Social motivation and implicit theory of mind in children with autism spectrum disorder. *Autism*

Research, *10*, 1834–1844. https://doi.org/10.1002/aur.1836

Bury, S. M., Jellett, R., Spoor, J. R., & Hedley, D. (2020). "It Defines Who I Am" or "It's Something I Have": What Language Do [Autistic] Australian Adults [on the Autism Spectrum] Prefer? *Journal of Autism and Developmental Disorders*, 1–11. https://doi.org/10.1007/s10803-020-04425-3

Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence and kappa. *Journal of Clinical Epidemiology*, *46*(5), 423–429. http://www.ncbi.nlm.nih.gov/pubmed/8501467

Cakir, J., Frye, R. E., & Walker, S. J. (2020). The lifetime social cost of autism: 1990–2029. *Research in Autism Spectrum Disorders*, *72*, 101502. https://doi.org/10.1016/j.rasd.2019.101502

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

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

Chen, T.-C., Hsieh, M. H., Lin, Y.-T., Chan, P.-Y. S., & Cheng, C.-H. (2020). Mismatch negativity to different deviant changes in autism spectrum disorders: A meta-analysis. *Clinical Neurophysiology*, *131*(3), 766–777. https://doi.org/10.1016/j.clinph.2019.10.031

Chen, Y.-T., Tsou, K.-S., Chen, H.-L., Wong, C.-C., Fan, Y.-T., & Wu, C.-T. (2018). Functional but Inefficient Kinesthetic Motor Imagery in Adolescents with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *48*(3), 784–795. https://doi.org/https://dx.doi.org/10.1007/s10803-017-3367-y

Chennu, S., Noreika, V., Gueorguiev, D., Blenkmann, A., Kochen, S., Ibáñez, A.,

Owen, A. M., & Bekinschtein, T. A. (2013). *Behavioral/Cognitive Expectation and Attention in Hierarchical Auditory Prediction*. https://doi.org/10.1523/JNEUROSCI.0114-13.2013

Cheung, M. W.-L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods*, *19*(2), 211–229. https://doi.org/10.1037/a0032968

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

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Academic Press.

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*.

Conner, C. M., Cramer, R. D., & McGonigle, J. J. (2019). Examining the Diagnostic Validity of Autism Measures Among Adults in an Outpatient Clinic Sample. *Autism in Adulthood*, *1*(1), 60–68. https://doi.org/10.1089/aut.2018.0023

Constant, A., Bervoets, J., Hens, K., & Van de Cruys, S. (2020). Precise Worlds for Certain Minds: An Ecological Perspective on the Relational Self in Autism. *Topoi*, *39*(3), 611–622. https://doi.org/10.1007/s11245-018-9546-4

Cook, J., Saygin, A. P., Swain, R., & Blakemore, S.-J. J. (2009). Reduced sensitivity to minimum-jerk biological motion in autism spectrum conditions. *Neuropsychologia*, *47*(14), 3275–3278. https://doi.org/10.1016/j.neuropsychologia.2009.07.010

Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, Not Autism, Predicts Poor Recognition of Emotional Facial Expressions: *Https://Doi.Org/10.1177/0956797612463582*, *24*(5), 723–732.

https://doi.org/10.1177/0956797612463582

Cook, R., Brewer, R., Shah, P., & Bird, G. (2014). Intact Facial Adaptation in Autistic Adults. *Autism Research*, *7*(4), 481–490. https://doi.org/10.1002/aur.1381

Couture, S. M., Penn, D. L., Losh, M., Adolphs, R., Hurley, R., & Piven, J. (2010). Comparison of social cognitive functioning in schizophrenia and high functioning autism: more convergence than divergence. *Psychological Medicine*, *40*(4), 569. https://doi.org/10.1017/S003329170999078X

Crawley, D., Zhang, L., Jones, E., Ahmad, J., San José Cáceres, A., Oakley, B., Charman, T., Buitelaar, J., Murphy, D., Chatham, C., den Ouden, H., Loth, E., & EU-AIMS LEAP group. (2019). Modelling cognitive flexibility in autism spectrum disorder and typical development reveals comparable developmental shifts in learning mechanisms Authors: *PsyArXiv*. https://doi.org/10.31234/osf.io/h7jcm

Cribb, S. J., Olaithe, M., Di Lorenzo, R., Dunlop, P. D., & Maybery, M. T. (2016). Embedded Figures Test Performance in the Broader Autism Phenotype: A Meta-analysis. *Journal of Autism and Developmental Disorders*, *46*(9), 2924–2939. https://doi.org/10.1007/s10803-016-2832-3

Cummings, K. K., Lawrence, K. E., Hernandez, L. M., Wood, E. T., Bookheimer, S. Y., Dapretto, M., & Green, S. A. (2020). Sex Differences in Salience Network Connectivity and its Relationship to Sensory Over-Responsivity in Youth with Autism Spectrum Disorder. *Autism Research*, *13*(9), 1489–1500. https://doi.org/10.1002/aur.2351

Cusack, J. P., Williams, J. H. G., & Neri, P. (2015). Action Perception Is Intact in Autism Spectrum Disorder. *The Journal of Neuroscienceeuroscience*, *35*(5), 1849–1857. https://doi.org/10.1523/JNEUROSCI.4133-13.2015

David, N., Rose, M., Schneider, T. R., Vogeley, K., & Engel, A. K. (2010). Brief report: Altered horizontal binding of single dots to coherent motion in autism. *Journal of Autism and Developmental Disorders*, *40*(12), 1549–1551.

https://doi.org/10.1007/s10803-010-1008-9

Davis, G., & Plaisted-Grant, K. (2015). Low endogenous neural noise in autism. *Autism*, *19*(3), 351–362. https://doi.org/10.1177/1362361314552198

De-Wit, L., Machilsen, B., & Putzeys, T. (2010). Predictive Coding and the Neural Response to Predictable Stimuli. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(26), 8702–8703. https://doi.org/10.1523/JNEUROSCI.2248-10.2010

de Gelder, B. (2009). Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1535), 3475–3484. https://doi.org/10.1098/rstb.2009.0190

Del Re, A. C. (2013). *compute.es: Compute Effect Sizes. R package* (0.2-2). http://cran.r-project.org/web/packages/compute.es

den Ouden, H. E. M., Daunizeau, J., Roiser, J., Friston, K. J., & Stephan, K. E. (2010). Striatal prediction error modulates cortical coupling. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(9), 3210–3219. https://doi.org/10.1523/JNEUROSCI.4458-09.2010

den Ouden, H. E. M., Friston, K. J., Daw, N. D., Mcintosh, A. R., & Stephan, K. E. (2009). A dual role for prediction error in associative learning. *Cerebral Cortex (New York, N.Y. : 1991)*, *19*(5), 1175–1185. https://doi.org/10.1093/cercor/bhn161

Downing, P. E. (2001). A Cortical Area Selective for Visual Processing of the Human Body. *Science*, *293*(5539), 2470–2473. https://doi.org/10.1126/science.1063414

Downing, P. E., & Peelen, M. V. (2011). The role of occipitotemporal body-selective regions in person perception. *Cognitive Neuroscience*, *2*(3–4), 186–203. https://doi.org/10.1080/17588928.2011.582945

Dumas, G., Soussignan, R., Hugueville, L., Martinerie, J., & Nadel, J. (2014). Revisiting mu suppression in autism spectrum disorder. *Brain Research*, *1585*, 108–119. https://doi.org/10.1016/j.brainres.2014.08.035

Dunn, M. A., Gomes, H., & Gravel, J. (2008). Mismatch Negativity in Children with Autism and Typical Development. *Journal of Autism and Developmental Disorders*, *38*(1), 52–71. https://doi.org/10.1007/s10803-007-0359-3

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463. http://www.ncbi.nlm.nih.gov/pubmed/10877304

Dyck, M. J., Ferguson, K., & Shochet, I. M. (2001). Do autism spectrum disorders differ from each other and from non-spectrum disorders on emotion recognition tests? *European Child & Adolescent Psychiatry*, *10*(2), 105–116. https://doi.org/10.1007/s007870170033

Edey, R., Cook, J., Brewer, R., Bird, G., & Press, C. (2019). Adults with autism spectrum disorder are sensitive to the kinematic features defining natural human motion. *Autism Research*, *12*(2), 284–294. https://doi.org/10.1002/aur.2052

Edwards, G., Vetter, P., McGruer, F., Petro, L. S., & Muckli, L. (2017). Predictive feedback to V1 dynamically updates with sensory input. *Scientific Reports*, *7*(1), 1–12. https://doi.org/10.1038/s41598-017-16093-y

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ (Clinical Research Ed.)*, *315*(7109), 629–634. http://www.ncbi.nlm.nih.gov/pubmed/9310563

Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. *NeuroImage*, *59*(3), 2349–2361. https://doi.org/10.1016/j.neuroimage.2011.09.017

Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., & Fox, P. T. (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Human Brain Mapping*, *30*(9), 2907–2926. https://doi.org/10.1002/hbm.20718

Ekman, M., Kok, P., & de Lange, F. P. (2017). Time-compressed preplay of anticipated events in human primary visual cortex. *Nature Communications*, *8*(1), 15276. https://doi.org/10.1038/ncomms15276

Falck-Ytter, T., Rehnberg, E., Bö Lte, S., & Bölte, S. (2013). Lack of Visual Orienting to Biological Motion and Audiovisual Synchrony in 3-Year-Olds with Autism. *PLoS ONE*, *8*(7), 3–7. https://doi.org/10.1371/journal.pone.0068816

Federici, A., Parma, V., Vicovaro, M., Radassao, L., Casartelli, L., & Ronconi, L. (2020). Anomalous Perception of Biological Motion in Autism: A Conceptual Review and Meta-Analysis. *Scientific Reports*, *10*(1), 4576. https://doi.org/10.1038/s41598-020-61252-3

Feldman, H., & Friston, K. J. (2010). Attention, Uncertainty, and Free-Energy. *Frontiers in Human Neuroscience*, *4*(December), 1–23. https://doi.org/10.3389/fnhum.2010.00215

Fletcher-Watson, S., & Happe, F. (2019). *Autism: A New Introduction to Psychological Theory and Current Debate*. Routledge.

Floris, D. L., Filho, J. O. A., Lai, M. C., Giavasis, S., Oldehinkel, M., Mennes, M., Charman, T., Tillmann, J., Dumas, G., Ecker, C., Dell'Acqua, F., Banaschewski, T., Moessnang, C., Baron-Cohen, S., Durston, S., Loth, E., Murphy, D. G. M., Buitelaar, J. K., Beckmann, C. F., … Di Martino, A. (2021). Towards robust and replicable sex differences in the intrinsic brain function of autism. *Molecular Autism*, *12*(1), 19. https://doi.org/10.1186/s13229-021-00415-z

Fombonne, E. (2018). Editorial: The rising prevalence of autism. *Journal of Child Psychology and Psychiatry*, *59*(7), 717–720.

https://doi.org/10.1111/jcpp.12941

Fox, N. A., Bakermans-Kranenburg, M. J., Yoo, K. H., Bowman, L. C., Cannon, E. N., Vanderwert, R. E., Ferrari, P. F., & van IJzendoorn, M. H. (2016). Assessing human mirror activity with EEG mu rhythm: A meta-analysis. *Psychological Bulletin*, *142*(3), 291–313. https://doi.org/10.1037/bul0000031

Fraiman, D., Saunier, G., Martins, E. F., & Vargas, C. D. (2014). Biological motion coding in the brain: Analysis of visually driven EEG functional networks. *PLoS ONE*, *9*(1). https://doi.org/10.1371/journal.pone.0084612

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

Fridenson-Hayo, S., Berggren, S., Lassalle, A., Tal, S., Pigat, D., Bölte, S., Baron-Cohen, S., & Golan, O. (2016). Basic and complex emotion recognition in children with autism: cross-cultural findings. *Molecular Autism*, *7*, 52. https://doi.org/10.1186/s13229-016-0113-9

Friston, K. J. (2008). Hierarchical models in the brain. *PLoS Computational Biology*, *4*(11). https://doi.org/10.1371/journal.pcbi.1000211

Friston, K. J. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences*, *13*(7), 293–301. https://doi.org/10.1016/j.tics.2009.04.005

Friston, K. J. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Friston, K. J., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active Inference: A Process Theory. *Neural Computation*, *29*(1), 1–49. https://doi.org/10.1162/NECO_a_00912

Friston, K. J., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology-Paris*, *100*, 70–87. https://doi.org/10.1016/j.jphysparis.2006.10.001

Friston, K. J., Lawson, R., & Frith, C. D. (2013). On hyperpriors and hypopriors: comment on Pellicano and Burr. *Trends in Cognitive Sciences*, *17*(1), 1. https://doi.org/10.1016/J.TICS.2012.11.003

Friston, K. J., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics*, *104*(1–2), 137–160. https://doi.org/10.1007/s00422-011-0424-z

Fujioka, T., Inohara, K., Okamoto, Y., Masuya, Y., Ishitobi, M., Saito, D. N., Jung, M., Arai, S., Matsumura, Y., Fujisawa, T. X., Narita, K., Suzuki, K., Tsuchiya, K. J., Mori, N., Katayama, T., Sato, M., Munesue, T., Okazawa, H., Tomoda, A., … Kosaka, H. (2016). Gazefinder as a clinical supplementary tool for discriminating between autism spectrum disorder and typical development in male adolescents and adults. *Molecular Autism*, *7*(1), 19. https://doi.org/10.1186/s13229-016-0083-y

Fujisawa, T. X., Tanaka, S., Saito, D. N., Kosaka, H., & Tomoda, A. (2014). Visual attention for social information and salivary oxytocin levels in preschool children with autism spectrum disorders: an eye-tracking study. *Frontiers in Neuroscience*, *8*, 295. https://doi.org/10.3389/fnins.2014.00295

Gargaro, B. A., May, T., Tonge, B. J., Sheppard, D. M., Bradshaw, J. L., & Rinehart, N. J. (2018). Attentional Mechanisms in Autism, ADHD, and Autism-ADHD Using a Local–Global Paradigm. *Journal of Attention Disorders*, *22*(14), 1320–1332. https://doi.org/10.1177/1087054715603197

Garrido, M. I., Rowe, E. G., Halász, V., & Mattingley, J. B. (2018). Bayesian Mapping Reveals That Attention Boosts Neural Responses to Predicted and Unpredicted Stimuli. *Cerebral Cortex*, *28*(5), 1771–1782. https://doi.org/10.1093/cercor/bhx087

Gervais, W. M., Reed, C. L., Beall, P. M., & Roberts, R. J. (2010). Implied body action directs spatial attention. *Attention, Perception, & Psychophysics*, *72*(6), 1437–1443. https://doi.org/10.3758/APP.72.6.1437

Ghanouni, P., Memari, A. H., Shayestehfar, M., Moshayedi, P., Gharibzadeh, S., & Ziaee, V. (2015). Biological Motion Perception Is Affected by Age and Cognitive Style in Children Aged 8–15. *Neurology Research International*, *2015*, 1–6. https://doi.org/10.1155/2015/594042

Gonzalez-Gadea, M. L., Chennu, S., Bekinschtein, T. A., Rattazzi, A., Beraudi, A., Tripicchio, P., Moyano, B., Soffita, Y., Steinberg, L., Adolfi, F., Sigman, M., Marino, J., Manes, F., & Ibanez, A. (2015). Predictive coding in autism spectrum disorder and attention deficit hyperactivity disorder. *Journal of Neurophysiology*, *114*(5), 2625–2636. https://doi.org/10.1152/jn.00543.2015

Gordon, N., Hohwy, J., Davidson, M. J., Boxtel, J. van, & Tsuchiya, N. (2019). *From Intermodulation Components to Perception and Cognition- a Review*. https://doi.org/https://doi.org/10.31219/osf.io/bwpt8

Gordon, N., Koenig-Robert, R., Tsuchiya, N., Van Boxtel, J. J. A., & Hohwy, J. (2017). Neural markers of predictive coding under perceptual uncertainty revealed with hierarchical frequency tagging. *ELife*, *6*. https://doi.org/10.7554/eLife.22749

Gordon, N., Tsuchiya, N., Koenig-Robert, R., & Hohwy, J. (2019). Expectation and attention increase the integration of top-down and bottom-up signals in perception through different pathways. *PLOS Biology*, *17*(4), e3000233. https://doi.org/10.1371/journal.pbio.3000233

Goris, J., Braem, S., Nijhof, A. D., Rigoni, D., Deschrijver, E., Van de Cruys, S., Wiersema, J. R., & Brass, M. (2018). Sensory Prediction Errors Are Less Modulated by Global Context in Autism Spectrum Disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(8), 667–674. https://doi.org/10.1016/J.BPSC.2018.02.003

Goris, J., Brass, M., Cambier, C., Delplanque, J., Wiersema, J. R., & Braem, S. (2019). The relation between preference for predictability and autistic traits. *Autism Research*, *00*, aur.2244. https://doi.org/10.1002/aur.2244

Goris, J., Silvetti, M., Verguts, T., Wiersema, J. R., Brass, M., & Braem, S. (2020). 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

Gough, D. (2007). Weight of Evidence: a framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education*, *22*(2), 213–228. https://doi.org/10.1080/02671520701296189

Grèzes, J., Wicker, B., Berthoz, S., & de Gelder, B. (2009). A failure to grasp the affective meaning of actions in autism spectrum disorder subjects. *Neuropsychologia*, *47*, 1816–1825. https://doi.org/10.1016/j.neuropsychologia.2009.02.021

Grosbras, M.-H., Beaton, S., & Eickhoff, S. B. (2012). Brain regions involved in human movement perception: A quantitative voxel-based meta-analysis. *Human Brain Mapping*, *33*(2), 431–454. https://doi.org/10.1002/hbm.21222

Hamilton, A. F. de C. (2013). Reflecting on the mirror neuron system in autism: A systematic review of current theories. *Developmental Cognitive Neuroscience*, *3*, 91–105. https://doi.org/10.1016/j.dcn.2012.09.008

Han, Z., Bi, Y., Chen, J., Chen, Q., He, Y., & Caramazza, A. (2013). Distinct regions of right temporal cortex are associated with biological and human-agent motion: functional magnetic resonance imaging and neuropsychological evidence. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(39), 15442–15453. https://doi.org/10.1523/JNEUROSCI.5868-12.2013

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. In *Journal of Autism and Developmental Disorders* (Vol. 36, Issue 1, pp. 5–25). Springer.

https://doi.org/10.1007/s10803-005-0039-0

Higgins, J. P. T., & Green, J. (2011). *Cochrane Handbook for Systematic Reviews of Interventions* (5.1.0). The Cochrane Collaboration. www.handbook.cochrane.org

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

Hirai, M., Gunji, A., Inoue, Y., Kita, Y., Hayashi, T., Nishimaki, K., Nakamura, M., Kakigi, R., & Inagaki, M. (2014). Differential electrophysiological responses to biological motion in children and adults with and without autism spectrum disorders. *Research in Autism Spectrum Disorders*, *8*(12), 1623–1634. https://doi.org/10.1016/j.rasd.2014.08.014

Hiris, E. (2007). Detection of biological and nonbiological motion. *Journal of Vision*, *7*(12). https://doi.org/10.1167/7.12.4

Hobson, H. M., & Bishop, D. V. M. (2016). Mu suppression – A good measure of the human mirror neuron system? *Cortex*, *82*, 290–310. https://doi.org/10.1016/j.cortex.2016.03.019

Hubert, B., Wicker, B., Moore, D. G., Monfardini, E., Duverger, H., Da Fonséca, D., Deruelle, C., Moore, G., Monfardini, E., Duverger, H., Da Fonseca, D., Derulle, C., Moore, D. G., Monfardini, E., Duverger, H., Da Fonséca, 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

Hudepohl, M. B., Robins, D. L., King, T. Z., & Henrich, C. C. (2015). The role of emotion perception in adaptive functioning of people with autism spectrum disorders. *Autism*, *19*(1), 107–112. https://doi.org/10.1177/1362361313512725

Ishida, R., Kamio, Y., & Nakamizo, S. (2009). Perceptual distortions of visual illusions in children with high-functioning autism spectrum disorder. *Psychologia: An International Journal of Psychological Sciences*, *52*(3), 175–187. https://doi.org/10.2117/psysoc.2009.175

Jack, A., Keifer, C. M., & Pelphrey, K. A. (2017). Cerebellar contributions to biological motion perception in autism and typical development. *Human Brain Mapping*, *38*(4), 1914–1932. https://doi.org/https://dx.doi.org/10.1002/hbm.23493

Jack, A., & Morris, J. P. (2014). Neocerebellar contributions to social perception in adolescents with autism spectrum disorder. *Developmental Cognitive Neuroscience*, *10*, 77–92. https://doi.org/10.1016/J.DCN.2014.08.001

Jensen, M. S., Yao, R., Street, W. N., & Simons, D. J. (2011). Change blindness and inattentional blindness. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(5), 529–546. https://doi.org/10.1002/wcs.130

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*(2), 201–211. https://link.springer.com/content/pdf/10.3758%2FBF03212378.pdf

Jones, C. R. G., Pickles, A., Falcaro, M., Marsden, A. J. S., Happé, F., Scott, S. K., Sauter, D., Tregay, J., Phillips, R. J., Baird, G., Simonoff, E., & Charman, T. (2011). A multimodal approach to emotion recognition ability in autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, *52*(3), 275–285. https://doi.org/10.1111/j.1469-7610.2010.02328.x

Jones, C. R. G., Swettenham, J., Charman, T., Marsden, A. J. S., Tregay, J., Baird, G., Simonoff, E., & Happé, F. (2011). No evidence for a fundamental visual motion processing deficit in adolescents with autism spectrum disorders. *Autism Research*, *4*(5), 347–357. https://doi.org/10.1002/aur.209

Kaiser, M. D., Hudac, C. M., Shultz, S., Lee, S. M., Cheung, C., Berken, A. M., Deen, B., Pitskel, N. B., Sugrue, D. R., Voos, A. C., Saulnier, C. A., Ventola, P., Wolf, J. M., Klin, A., Vander Wyk, B. C., & Pelphrey, K. A. (2010). Neural

signatures of autism. *PNAS*, *107*(49), 21223–21228. https://doi.org/10.1073/pnas.1010412107

Kaiser, M. D., & Pelphrey, K. A. (2012). Disrupted action perception in autism: Behavioral evidence, neuroendophenotypes, and diagnostic utility. *Developmental Cognitive Neuroscience*, *2*(1), 25–35. https://doi.org/10.1016/J.DCN.2011.05.005

Kaiser, M. D., & Shiffrar, M. (2009). The visual perception of motion by observers with autism spectrum disorders: a review and synthesis. *Psychonomic Bulletin & Review*, *16*(5), 761–777. https://doi.org/10.3758/PBR.16.5.761

Kalandadze, T., Norbury, C., Nærland, T., & Næss, K. A. B. (2018). Figurative language comprehension in individuals with autism spectrum disorder: A meta-analytic review. In *Autism* (Vol. 22, Issue 2, pp. 99–117). SAGE Publications Ltd. https://doi.org/10.1177/1362361316668652

Karaminis, T., Arrighi, R., Forth, G., Burr, D., & Pellicano, E. (2020). Adaptation to the Speed of Biological Motion in Autism. *Journal of Autism and Developmental Disorders*, *50*(2), 373–385. https://doi.org/10.1007/s10803-019-04241-4

Karuppali, S. (2018). Do children with autism spectrum disorders exhibit biological motion perception deficits? Evidence using an action recognition paradigm. *Journal of Indian Association of Child and Adolescent Mental Health*, *14*(4), 38–57. http://www.jiacam.org/1404/orig2Oct2018.pdf

Karvelis, P., Seitz, A. R., Lawrie, S. M., & Seriès, P. (2018). Autistic traits, but not schizotypy, predict increased weighting of sensory information in Bayesian visual integration. *ELife*, *7*, e34115. https://doi.org/https://doi.org/10.7554/eLife.34115.001

Keehn, B., Müller, R.-A., & Townsend, J. (2013). Atypical attentional networks and the emergence of autism. *Neuroscience & Biobehavioral Reviews*, *37*(2), 164–183. https://doi.org/10.1016/j.neubiorev.2012.11.014

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism*, *20*(4), 442–462. https://doi.org/10.1177/1362361315588200

Kim, Y. S., Leventhal, B. L., Koh, Y.-J., Fombonne, E., Laska, E., Lim, E.-C., Cheon, K.-A., Kim, S.-J., Kim, Y.-K., Lee, H., Song, D.-H., & Grinker, R. R. (2011). Prevalence of Autism Spectrum Disorders in a Total Population Sample. *American Journal of Psychiatry*, *168*(9), 904–912. https://doi.org/10.1176/appi.ajp.2011.10101532

Klein, R. M. (2000). Inhibition of return. In *Trends in Cognitive Sciences* (Vol. 4, Issue 4, pp. 138–147). Elsevier Current Trends. https://doi.org/10.1016/S1364-6613(00)01452-2

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, *36*(ECVP Abstract Suppl), 14.

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

Kmet, L. M., Lee, R. C., & Cook, L. S. (2004). *Standard quality assessment criteria for evaluating primary research papers from a variety of fields*. University of Alberta Libraries. https://doi.org/10.7939/R37M04F16

Koenig-Robert, R., & VanRullen, R. (2013). SWIFT: A novel method to track the neural correlates of recognition. *NeuroImage*, *81*, 273–282. https://doi.org/10.1016/j.neuroimage.2013.04.116

Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex. *Neuron*, *75*(2), 265–270. https://doi.org/10.1016/J.NEURON.2012.04.034

Kok, P., Rahnev, D., Jehee, J. F. M., Lau, H. C., & de Lange, F. P. (2012). Attention Reverses the Effect of Prediction in Silencing Sensory Signals.

Cerebral Cortex, *22*(9), 2197–2206. https://doi.org/10.1093/cercor/bhr310

Koldewyn, K., Jiang, Y. V., Weigelt, S., & Kanwisher, N. (2013). Global/local processing in autism: not a disability, but a disinclination. *Journal of Autism and Developmental Disorders*, *43*(10), 2329–2340. https://doi.org/10.1007/s10803-013-1777-z

Koldewyn, K., Whitney, D., & Rivera, S. M. (2010). The psychophysics of visual motion and global form processing in autism. *Brain*, *133*(2), 599–610. https://doi.org/10.1093/brain/awp272

Koldewyn, K., Whitney, D., & Rivera, S. M. (2011). Neural correlates of coherent and biological motion perception in autism. *Developmental Science*, *14*(5), 1075–1088. https://doi.org/10.1111/j.1467-7687.2011.01058.x

Krakowski, A. (2014). Biological motion processing in typical development and in the autism spectrum [City University of New York]. In *City University of New York (CUNY)*. http://academicworks.cuny.edu/gc_etds

Kröger, A., Bletsch, A., Krick, C., Siniatchkin, M., Jarczok, T. A., Freitag, C. M., & Bender, S. (2014). Visual event-related potentials to biological motion stimuli in autism spectrum disorders. *Social Cognitive and Affective Neuroscience*, *9*(8), 1214–1222. https://doi.org/10.1093/scan/nst103

Krüger, B., Kaletsch, M., Pilgramm, S., Schwippert, S., Hennig, J., Stark, R., Lis, S., Gallhofer, B., Sammer, G., Zentgraf, K., & Munzert, J. (2017). Perceived Intensity of Emotional Point-Light Displays is Reduced in Subjects with ASD. *Journal of Autism and Developmental Disorders*, *48*(1). https://doi.org/10.1007/s10803-017-3286-y

Lancaster, J. L., & Martinez, M. J. (n.d.). *Mango: Multi-image Analysis GUI* (v4.0.1). http://rii.uthscsa.edu/mango/

Landrigan, P. J. (2010). What causes autism? Exploring the environmental contribution. *Current Opinion in Pediatrics*, *22*(2), 219–225. https://doi.org/10.1097/MOP.0b013e328336eb9a

Lawrence, K. E., Hernandez, L. M., Bowman, H. C., Padgaonkar, N. T., Fuster, E., Jack, A., Aylward, E., Gaab, N., Van Horn, J. D., Bernier, R. A., Geschwind, D. H., McPartland, J. C., Nelson, C. A., Webb, S. J., Pelphrey, K. A., Green, S. A., Bookheimer, S. Y., & Dapretto, M. (2020). Sex Differences in Functional Connectivity of the Salience, Default Mode, and Central Executive Networks in Youth with ASD. *Cerebral Cortex*, *30*(9), 5107–5120. https://doi.org/10.1093/cercor/bhaa105

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*(May), 302. https://doi.org/10.3389/fnhum.2014.00302

Lenth, R. V. (2021). *emmeans: Estimated Marginal Means, aka Least-Squares Means* (1.6.1). https://cran.r-project.org/package=emmeans

Lestou, V., Pollick, F. E., & Kourtzi, Z. (2008). Neural Substrates for Action Understanding at Different Description Levels in the Human Brain. *Journal of Cognitive Neuroscience*, *20*(2), 324–341. https://doi.org/10.1162/jocn.2008.20021

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

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*(3), 205–223. http://www.ncbi.nlm.nih.gov/pubmed/11055457

Manning, C., Kilner, J., Neil, L., Karaminis, T., & Pellicano, E. (2017). Children on the autism spectrum update their behaviour in response to a volatile environment. *Developmental Science*, *20*(5), e12435. https://doi.org/10.1111/desc.12435

Marsh, L. E., & Hamilton, A. F. de C. (2011). Dissociation of mirroring and mentalising systems in autism. *NeuroImage*, *56*(3), 1511–1519. https://doi.org/10.1016/J.NEUROIMAGE.2011.02.003

Masi, A., DeMayo, M. M., Glozier, N., & Guastella, A. J. (2017). An Overview of Autism Spectrum Disorder, Heterogeneity and Treatment Options. *Neuroscience Bulletin*, *33*(2), 183–193. https://doi.org/10.1007/s12264-017-0100-y

MathWorks. (2019). *MATLAB* (No. R2019b). The MathWorks Inc.

Matson, J. L., & Neal, D. (2009). Diagnosing high incidence autism spectrum disorders in adults. *Research in Autism Spectrum Disorders*, *3*(3), 581–589. https://doi.org/10.1016/j.rasd.2009.01.001

Maunsell, J. H. R. (2015). *Neuronal Mechanisms of Visual Attention*. https://doi.org/10.1146/annurev-vision-082114-035431

McClelland, G. H., John G. Lynch, Irwin, J. R., Spiller, S. A., & Fitzsimons, G. J. (2015). Median splits, Type II errors, and false–positive consumer psychology: Don't fight the power. *Journal of Consumer Psychology*, *25*(4), 679–689. https://doi.org/10.1016/J.JCPS.2015.05.006

McFayden, T. C., Antezana, L., Albright, J., Muskett, A., & Scarpa, A. (2020). Sex Differences in an Autism Spectrum Disorder Diagnosis: Are Restricted Repetitive Behaviors and Interests the Key? *Review Journal of Autism and Developmental Disorders*, *7*(2), 119–126. https://doi.org/10.1007/s40489-019-00183-w

McKay, L. S., Simmons, D. R., McAleer, P., Marjoram, D., Piggot, J., & Pollick, F. E. (2012). Do distinct atypical cortical networks process biological motion

information in adults with Autism Spectrum Disorders? *NeuroImage*, *59*(2), 1524–1533. https://doi.org/10.1016/j.neuroimage.2011.08.033

Mehrpour, V., Martinez-Trujillo, J. C., & Treue, S. (2020). Attention amplifies neural representations of changes in sensory input at the expense of perceptual accuracy. *Nature Communications*, *11*(1), 2128. https://doi.org/10.1038/s41467-020-15989-0

Meilleur, A.-A. S., Jelenic, P., & Mottron, L. (2014). Prevalence of Clinically and Empirically Defined Talents and Strengths in Autism. *Journal of Autism and Developmental Disorders*, 1354–1367. https://doi.org/10.1007/s10803-014-2296-2

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, *4*(1), 1. https://doi.org/10.1186/2046-4053-4-1

Morrison, K. E., Pinkham, A. E., Kelsven, S., Ludwig, K., Penn, D. L., & Sasson, N. J. (2019). Psychometric Evaluation of Social Cognitive Measures for Adults with Autism. *Autism Research*, *12*(5), 766–778. https://doi.org/10.1002/aur.2084

Mottron, L., Bouvet, L., Bonnel, A., Samson, F., Burack, J. a., Dawson, M., & Heaton, P. (2013). Veridical mapping in the development of exceptional autistic abilities. *Neuroscience and Biobehavioral Reviews*, *37*(2), 209–228. https://doi.org/10.1016/j.neubiorev.2012.11.016

Mottron, L., Dawson, M., Soulières, I., Hubert, B., & Burack, J. (2006). Enhanced Perceptual Functioning in Autism: An Update, and Eight Principles of Autistic Perception. *Journal of Autism and Developmental Disorders 2006*

36:1, *36*(1), 27–43. https://doi.org/10.1007/s10803-005-0040-7

Muckli, L., De Martino, F., Vizioli, L., Petro, L. S., Smith, F. W., Ugurbil, K., Goebel, R., & Yacoub, E. (2015). Contextual Feedback to Superficial Layers of V1. *Current Biology*, *25*(20), 2690–2695. https://doi.org/10.1016/j.cub.2015.08.057

Muler, E., Schuler, A., & Yates, G. B. (2008). Social challenges and supports from the perspective of individuals with Asperger syndrome and other autism spectrum disabilities. *Autism*, *12*(2), 173–190. https://doi.org/10.1177/1362361307086664

Murphy, G., Groeger, J. A., & Greene, C. M. (2016). Twenty years of load theory—Where are we now, and where should we go next? *Psychonomic Bulletin and Review*, *23*(5), 1316–1340. https://doi.org/10.3758/s13423-015-0982-5

Murphy, P., Brady, N., Fitzgerald, M., & Troje, N. F. (2009). No evidence for impaired perception of biological motion in adults with autistic spectrum disorders. *Neuropsychologia*, *47*(14), 3225–3235. https://doi.org/10.1016/j.neuropsychologia.2009.07.026

Nackaerts, E., Wagemans, J., Helsen, W., Swinnen, S. P., Wenderoth, N., Alaerts, K., Schmitz, C., Hospitalier, C., & Vinatier, L. (2012). Recognizing Biological Motion and Emotions from Point-Light Displays in Autism Spectrum Disorders. *PLoS ONE*, *7*(9), e44473. https://doi.org/10.1371/journal.pone.0044473

Neil, L., Olsson, N. C., & Pellicano, E. (2016). The Relationship Between Intolerance of Uncertainty, Sensory Sensitivities, and Anxiety in Autistic and Typically Developing Children. *Journal of Autism and Developmental Disorders*, *46*(6), 1962–1973. https://doi.org/10.1007/s10803-016-2721-9

Neri, P., Morrone, M. C., & Burr, D. C. (1998). Seeing biological motion. *Nature*, *395*(6705), 894–896. https://doi.org/10.1038/27661

Noble, K., Glowinski, D., Murphy, H., Jola, C., McAleer, P., Darshane, N., Penfield, K., Kalyanasundaram, S., Camurri, A., & Pollick, F. E. (2014). Event Segmentation and Biological Motion Perception in Watching Dance. *Art & Perception*, *2*(1–2), 59–74. https://doi.org/10.1163/22134913-00002011

Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottereau, B. R., & Rossion, B. (2015). The steady-state visual evoked potential in vision research: A review. *Journal of Vision*, *15*(6), 4. https://doi.org/10.1167/15.6.4

O'Connor, K. (2012). Auditory processing in autism spectrum disorder: A review. *Neuroscience & Biobehavioral Reviews*, *36*(2), 836–854. https://doi.org/10.1016/j.neubiorev.2011.11.008

Oberman, L. M., Hubbard, E. M., McCleery, J. P., Altschuler, E. L., Ramachandran, V. S., & Pineda, J. A. (2005). EEG evidence for mirror neuron dysfunction in autism spectrum disorders. *Cognitive Brain Research*, *24*(2), 190–198. https://doi.org/10.1016/j.cogbrainres.2005.01.014

Oberman, L. M., McCleery, J. P., Hubbard, E. M., Bernier, R., Wiersema, J. R., Raymaekers, R., & Pineda, J. A. (2013). Developmental changes in mu suppression to observed and executed actions in autism spectrum disorders. *Social Cognitive and Affective Neuroscience*, *8*(3), 300–304. https://doi.org/10.1093/scan/nsr097

Oberman, L. M., Ramachandran, V. S., & Pineda, J. A. (2008). Modulation of mu suppression in children with autism spectrum disorders in response to familiar or unfamiliar stimuli: The mirror neuron hypothesis. *Neuropsychologia*, *46*(5), 1558–1565. https://doi.org/10.1016/j.neuropsychologia.2008.01.010

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Orekhova, E. V., & Stroganova, T. A. (2014). Arousal and attention re-orienting

in autism spectrum disorders: evidence from auditory event-related potentials. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00034

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210. https://doi.org/10.1186/s13643-016-0384-4

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

Palmer, C. J., Paton, B., Kirkovski, M., Enticott, P., & Hohwy, J. (2015). Autism and the sensorimotor effects of the Rubber-Hand Illusion. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/conf.fnhum.2015.217.00227

Papagiannopoulou, E. A., Chitty, K. M., Hermens, D. F., Hickie, I. B., & Lagopoulos, J. (2014). A systematic review and meta-analysis of eye-tracking studies in children with autism spectrum disorders. *Social Neuroscience*, *9*(6), 610–632. https://doi.org/10.1080/17470919.2014.934966

Parr, T., & Friston, K. J. (2017). Uncertainty, epistemics and active Inference. *Journal of the Royal Society Interface*, *14*(136). https://doi.org/10.1098/rsif.2017.0376

Parr, T., & Friston, K. J. (2019). Attention or salience? *Current Opinion in Psychology*, *29*, 1–5. https://doi.org/10.1016/j.copsyc.2018.10.006

Parron, C., De Fonseca, D., Santos, A., Monfardini, E., Deruelle, C., Da Fonseca, D., Santos, A., Moore, D. G., Monfardini, E., & Deruelle, C. (2008). Recognition of biological motion in children with autistic spectrum disorders. *Autism*, *12*(3), 261–274. https://doi.org/10.1177/1362361307089520

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

Peelen, M. V., Atkinson, A. P., Andersson, F., & Vuilleumier, P. (2007). Emotional modulation of body-selective visual areas. *Social Cognitive and Affective Neuroscience*, *2*(4), 274–283. https://doi.org/10.1093/scan/nsm023

Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nature Reviews Neuroscience*, *8*(8), 636–648. https://doi.org/10.1038/nrn2195

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pell, P. J., Mareschal, I., Calder, A. J., Von Dem Hagen, E. A. H., Clifford, C. W. G., Baron-Cohen, S., & Ewbank, M. P. (2016). Intact priors for gaze direction in adults with high-functioning autism spectrum conditions. *Molecular Autism*, *7*(1), 25. https://doi.org/10.1186/s13229-016-0085-9

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897X00366

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

Pellicano, E., Gibson, L., Maybery, M., Durkin, K., & Badcock, D. R. (2005). Abnormal global processing along the dorsal visual pathway in autism: A possible mechanism for weak visuospatial coherence? *Neuropsychologia*, *43*(7), 1044–1053. https://doi.org/10.1016/j.neuropsychologia.2004.10.003

Pellicano, E., Jeffery, L., Burr, D., & Rhodes, G. (2007). Abnormal Adaptive Face-Coding Mechanisms in Children with Autism Spectrum Disorder. *Current Biology*, *17*(17), 1508–1512. https://doi.org/10.1016/J.CUB.2007.07.065

Pelphrey, K. A., Shultz, S., Hudac, C. M., & Vander Wyk, B. C. (2011). Research Review: Constraining heterogeneity: the social brain and its development in autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, *52*(6), 631–644. https://doi.org/10.1111/j.1469-7610.2010.02349.x

Philip, R. C. M., Dauvermann, M. R., Whalley, H. C., Baynham, K., Lawrie, S. M., & Stanfield, A. C. (2012). A systematic review and meta-analysis of the fMRI investigation of autism spectrum disorders. *Neuroscience and Biobehavioral Reviews*, *36*(2), 901–942. https://doi.org/10.1016/j.neubiorev.2011.10.008

Philip, R. C. M., Whalley, H. C., Stanfield, a C., Sprengelmeyer, R., Santos, I. M., Young, a W., Atkinson, A. P., Calder, A. J., Johnstone, E. C., Lawrie, S. M., & Hall, J. (2010). Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychological Medicine*, *40*(11), 1919–1929. https://doi.org/10.1017/S0033291709992364

Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition*, *82*, B51–B61. https://ac.els-cdn.com/S0010027701001470/1-s2.0-S0010027701001470-main.pdf?_tid=4a2133d8-bfc6-11e7-be7e-00000aab0f27&acdnat=1509624582_1292ab797a880b6fff875b7441c1c349

Popay, J., Roberts, H., Sowden, A., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006). *Guidance on the conduct of narrative synthesis in systematic reviews: A product from the ESRC Methods Programme*. Lancaster University. https://doi.org/10.13140/2.1.1018.4643

Press, C., Kok, P., & Yon, D. (2020). The Perceptual Prediction Paradox. *Trends in Cognitive Sciences*, *24*(1), 13–24. https://doi.org/10.1016/j.tics.2019.11.003

Press, C., & Yon, D. (2019). Perceptual Prediction: Rapidly Making Sense of a Noisy World. *Current Biology*, *29*(15), R751–R753. https://doi.org/10.1016/J.CUB.2019.06.054

Price, K. J., Shiffrar, M., & Kerns, K. A. (2011). Movement perception and movement production in Asperger's Syndrome Previously researchers have documented movement impairments in individuals with Asperger's Syndrome (AS) and autism (e.g. *Research in Autism Spectrum Disorders*, *6*, 391–398. https://doi.org/10.1016/j.rasd.2011.06.013

Price, K. J., Shiffrar, M., & Kerns, K. A. (2012). Movement perception and movement production in Asperger's Syndrome. *Research in Autism Spectrum Disorders*, *6*(1), 391–398. https://doi.org/10.1016/j.rasd.2011.06.013

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

R Core Team. (2020). *R: A language and environment for statistical computing* (4.0.1). R Foundation for Statistical Computing. https://www.r-project.org/

R Core Team. (2021). *R: A language and environment for statistical computing*. (4.0.4). R Foundation for Statistical Computing. https://www.r-project.org/

Raymaekers, R., Wiersema, J. R., & Roeyers, H. (2009). EEG study of the mirror neuron system in children with high functioning autism. *Brain Research*, *1304*, 113–121. https://doi.org/10.1016/j.brainres.2009.09.068

Remington, A. M., Swettenham, J. G., & Lavie, N. (2012). Lightening the Load: Perceptual Load Impairs Visual Detection in Typical Adults but Not in Autism. *Journal of Abnormal Psychology*, *121*(2), 544–551. https://doi.org/10.1037/a0027670

Rendina-Gobioff, G., & Kromery, J. D. (2006). PUB_BIAS: a SAS macro for detecting publication bias in meta-analysis. *14th Annual SouthEast SAS Users Group (SESUG) Conference*. https://analytics.ncsu.edu/sesug/2006/PO04_06.PDF

Richard, F. D., Bond, C. F., & Stokes-Zoota, J. J. (2003). One Hundred Years of Social Psychology Quantitatively Described. *Review of General Psychology*, *7*(4), 331–363. https://doi.org/10.1037/1089-2680.7.4.331

Robic, S., Sonié, S., Fonlupt, P., Henaff, M. A., Touil, N., Coricelli, G., Mattout, J., & Schmitz, C. (2015). Decision-Making in a Changing World: A Study in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *45*(6), 1603–1613. https://doi.org/10.1007/s10803-014-2311-7

Rogge, N., & Janssen, J. (2019). The Economic Costs of Autism Spectrum Disorder: A Literature Review. *Journal of Autism and Developmental Disorders*, *49*(7), 2873–2900. https://doi.org/10.1007/s10803-019-04014-z

RStudio Team. (2016). *RStudio: Integrated Development for R*. RStudio INC. http://www.rstudio.com/

RStudio Team. (2020). *RStudio: Integrated Development Environment for R* (1.3.1093). RStudio, PBC. http://www.rstudio.com/

Russell-Smith, S. N., Bayliss, D. M., Maybery, M. T., & Tomkinson, R. L. (2013). Are the autism and positive schizotypy spectra diametrically opposed in empathizing and systemizing? *Journal of Autism and Developmental Disorders*, *43*(3), 695–706. https://doi.org/10.1007/s10803-012-1614-9

Rutherford, M. D., & Troje, N. F. (2012). IQ predicts biological motion perception in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *42*(4), 557–565. https://doi.org/10.1007/s10803-011-1267-0

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*(1), 2. https://doi.org/10.1186/2040-2392-6-2

Sajid, N., Ball, P. J., & Friston, K. J. (2019). *Active inference: demystified and*

compared. http://arxiv.org/abs/1909.10863

Sanders, L. L. O., Muckli, L., de Millas, W., Lautenschlager, M., Heinz, A., Kathmann, N., & Sterzer, P. (2012). Detection of visual events along the apparent motion trace in patients with paranoid schizophrenia. *Psychiatry Research*, *198*(2), 216–223. https://doi.org/10.1016/j.psychres.2012.03.006

SAS OnDemand for Academics: SAS Studio (Release 3.7). (n.d.). SAS Institute Inc.

Saygin, A. P., Cook, J., & Blakemore, S.-J. J. (2010). Unaffected perceptual thresholds for biological and non-biological form-from-motion perception in autism spectrum conditions. *PLoS ONE*, *5*(10), 1–7. https://doi.org/10.1371/journal.pone.0013491

Schaalje, G. B., McBride, J. B., & Fellingham, G. W. (2001). Approximations to distributions of test statistics in complex mixed linear models using SAS Proc MIXED. *SAS Conference Proceedings: SAS Users Group International 26*, 262. http://www2.sas.com/proceedings/sugi26/p262-26.pdf

Scheepers, C. (2014). *Between-group matching of confounding variables: Why covariates remain important for analysis*. https://doi.org/10.13140/2.1.3623.8726

Schwartz, S., Shinn-Cunningham, B., & Tager-Flusberg, H. (2018). Meta-analysis and systematic review of the literature characterizing auditory mismatch negativity in individuals with autism. *Neuroscience & Biobehavioral Reviews*, *87*, 106–117. https://doi.org/10.1016/j.neubiorev.2018.01.008

Schwiedrzik, C. M., Alink, A., Kohler, A., Singer, W., & Muckli, L. (2007). A spatio-temporal interaction on the apparent motion trace. *Vision Research*, *47*(28), 3424–3433. https://doi.org/10.1016/j.visres.2007.10.004

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

Shi, J., Weng, X., He, S., & Jiang, Y. (2010). Biological motion cues trigger reflexive attentional orienting. *Cognition*, *117*(3), 348–354. https://doi.org/10.1016/j.cognition.2010.09.001

Simmons, D. R., McKay, L., McAleer, P., Toal, E., Robertson, A., & Pollick, F. E. (2007). Neural noise and autism spectrum disorders. *Perception*, *36*, 119–120.

Simmons, D. R., & Milne, E. (2015). Response to Davis and Plaisted-Grant: Low or high endogenous neural noise in autism spectrum disorder? *Autism*, *19*(3), 363–364. https://doi.org/10.1177/1362361314557683

Simmons, D. R., Robertson, A. E., McKay, L. S., Toal, E., McAleer, P., & Pollick, F. E. (2009). Vision in autism spectrum disorders. *Vision Research*, *49*(22), 2705–2739. https://doi.org/10.1016/j.visres.2009.08.005

Simmons, D. R., & Todorova, G. K. (2018). Local Versus Global Processing in Autism: Special Section Editorial. *Journal of Autism and Developmental Disorders*, *48*(4), 1338–1340. https://doi.org/10.1007/s10803-017-3452-2

Skripkauskaite, S., Slade, L., & Mayer, J. (2021). Attentional shifting differences in autism: Domain general, domain specific or both? *Autism*. https://doi.org/10.1177/13623613211001619

Smith, R. E. W., Avery, J. A., Wallace, G. L., Kenworthy, L., Gotts, S. J., & Martin, A. (2019). Sex differences in resting-state functional connectivity of the cerebellum in autism spectrum disorder. *Frontiers in Human Neuroscience*, *13*, 104. https://doi.org/10.3389/fnhum.2019.00104

Sosnowy, C., Silverman, C., Shattuck, P., & Garfield, T. (2019). Setbacks and Successes: How Young Adults on the Autism Spectrum Seek Friendship. *Autism in Adulthood*, *1*(1), 44–51. https://doi.org/10.1089/aut.2018.0009

Sotoodeh, M. S., Taheri-Torbati, H., Sohrabi, M., & Ghoshuni, M. (2019). Perception of biological motions is preserved in people with autism spectrum disorder: electrophysiological and behavioural evidences. *Journal*

of Intellectual Disability Research, *63*(1), 72–84. https://doi.org/10.1111/jir.12565

Southwell, R., Baumann, A., Gal, C., Barascud, N., Friston, K., & Chait, M. (2017). Is predictability salient? A study of attentional capture by auditory patterns. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1714). https://doi.org/10.1098/rstb.2016.0105

Southwell, R., & Chait, M. (2018). Enhanced deviant responses in patterned relative to random sound sequences. *Cortex*, *109*, 92–103. https://doi.org/10.1016/j.cortex.2018.08.032

Sparrow, W. A., Shinkfield, A. J., Day, R. H., & Zerman, L. (1999). Visual Perception of Human Activity and Gender in Biological-Motion Displays by Individuals With Mental Retardation. *American Journal on Mental Retardation*, *104*(3), 215. https://doi.org/10.1352/0895-8017(1999)104<0215:VPOHAA>2.0.CO;2

St John-Saaltink, E., Utzerath, C., Kok, P., Lau, H. C., & de Lange, F. P. (2015). Expectation Suppression in Early Visual Cortex Depends on Task Set. *PloS One*, *10*(6), e0131172. https://doi.org/10.1371/journal.pone.0131172

Sterne, J. A. C., & Egger, M. (2001). Funnel plots for detecting bias in meta-analysis. *Journal of Clinical Epidemiology*, *54*(10), 1046–1055. https://doi.org/10.1016/S0895-4356(01)00377-8

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rucker, G., Harbord, R. M., Schmid, C. H., Tetzlaff, J., Deeks, J. J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D. G., Moher, D., & Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*, *343*(jul22 1), d4002–d4002. https://doi.org/10.1136/bmj.d4002

Sterzer, P., Adams, R. A., Fletcher, P., Frith, C., Lawrie, S. M., Muckli, L., Petrovic, P., Uhlhaas, P., Voss, M., & Corlett, P. R. (2018). The Predictive

Coding Account of Psychosis. *Biological Psychiatry*, *0*(0). https://doi.org/10.1016/j.biopsych.2018.05.015

Stewart, M. E., Watson, J., Allcock, A.-J., & Yaqoob, T. (2009). Autistic traits predict performance on the block design. *Autism*, *13*(2), 133–142. https://doi.org/10.1177/1362361308098515

Sun, X., Allison, C., Wei, L., Matthews, F. E., Auyeung, B., Wu, Y. Y., Griffiths, S., Zhang, J., Baron-Cohen, S., & Brayne, C. (2019). Autism prevalence in China is comparable to Western prevalence. *Molecular Autism*, *10*(1), 7. https://doi.org/10.1186/s13229-018-0246-0

Swettenham, J., Remington, A., Laing, K., Fletcher, R., Coleman, M., & Gomez, J.-C. C. (2013). Perception of Pointing from Biological Motion Point-Light Displays in Typically Developing Children and Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *43*(6), 1437–1446. https://doi.org/10.1007/s10803-012-1699-1

Szatmari, P., Georgiades, S., Bryson, S., Zwaigenbaum, L., Roberts, W., Mahoney, W., Goldberg, J., & Tuff, L. (2006). Investigating the structure of the restricted, repetitive behaviours and interests domain of autism. *Journal OfChild Psychology and Psychiatry*, *47*(6), 582–590. https://doi.org/10.1111/j.1469-7610.2005.01537.x

Thompson, J. C., & Baccus, W. (2012). Form and motion make independent contributions to the response to biological motion in occipitotemporal cortex. *NeuroImage*, *59*(1), 625–634. https://doi.org/10.1016/j.neuroimage.2011.07.051

Thornton, I. M., & Vuong, Q. C. (2004). Incidental Processing of Biological Motion. *Current Biology*, *14*(12), 1084–1089. https://doi.org/10.1016/j.cub.2004.06.025

Thurman, S. M., van Boxtel, J. J. A., Monti, M. M., Chiang, J. N., & Lu, H. (2016). Neural adaptation in pSTS correlates with perceptual aftereffects to biological motion and with autistic traits. *NeuroImage*, *136*, 149–161.

https://doi.org/10.1016/j.neuroimage.2016.05.015

Todd, J., & Cornwell, R. (2018). The importance of precision to updating models of the sensory environment. *Biological Psychology*, *139*, 8–16. https://doi.org/10.1016/j.biopsycho.2018.09.012

Todorova, G. K., Hatton, R. E. M., & Pollick, F. E. (2019). Biological motion perception in autism spectrum disorder: a meta-analysis. *Molecular Autism*, *10*(1), 49. https://doi.org/10.1186/s13229-019-0299-8

Todorovic, A., & de Lange, F. P. (2012). Repetition Suppression and Expectation Suppression Are Dissociable in Time in Early Auditory Evoked Fields. *Journal of Neuroscience*, *32*(39), 13389–13395. https://doi.org/10.1523/JNEUROSCI.2227-12.2012

Treviño, M., De la Torre-Valdovinos, B., & Manjarrez, E. (2016). Noise Improves Visual Motion Discrimination via a Stochastic Resonance-Like Phenomenon. *Frontiers in Human Neuroscience*, *10*(NOV2016), 572. https://doi.org/10.3389/fnhum.2016.00572

Tulver, K., Aru, J., Rutiku, R., & Bachmann, T. (2019). Individual differences in the effects of priors on perception: A multi-paradigm approach. *Cognition*, *187*, 167–177. https://doi.org/10.1016/j.cognition.2019.03.008

Turi, M., Burr, D. C., Igliozzi, R., Aagten-Murphy, D., Muratori, F., & Pellicano, E. (2015). Children with autism spectrum disorder show reduced adaptation to number. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(25), 7868–7872. https://doi.org/10.1073/pnas.1504099112

Turi, M., Muratori, F., Tinelli, F., Morrone, M. C., & Burr, D. C. (2017). Autism is associated with reduced ability to interpret grasping actions of others. *Scientific Reports*, *7*, 12687. https://doi.org/10.1038/s41598-017-12995-z

Turkeltaub, P. E., Eickhoff, S. B., Laird, A. R., Fox, M., Wiener, M., & Fox, P. (2012). Minimizing within-experiment and within-group effects in activation

likelihood estimation meta-analyses. *Human Brain Mapping*, *33*(1), 1–13. https://doi.org/10.1002/hbm.21186

Tye, C., Runicles, A. K., Whitehouse, A. J. O., & Alvares, G. A. (2019). Characterizing the Interplay Between Autism Spectrum Disorder and Comorbid Medical Conditions: An Integrative Review. *Frontiers in Psychiatry*, *9*. https://doi.org/10.3389/fpsyt.2018.00751

Ulloa, E. R., & Pineda, J. A. (2007). Recognition of point-light biological motion: Mu rhythms and mirror neuron activity. *Behavioural Brain Research*, *183*(2), 188–194. https://doi.org/10.1016/j.bbr.2007.06.007

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How Many Studies Do You Need? *Journal of Educational and Behavioral Statistics*, *35*(2), 215–247. https://doi.org/10.3102/1076998609346961

Van Boxtel, J. J. A., Dapretto, M., & Lu, H. (2016). Intact recognition, but attenuated adaptation, for biological motion in youth with autism spectrum disorder. *Autism Research*, *9*(10), 1103–1113. https://doi.org/10.1002/aur.1595

van Boxtel, J. J. A., & Lu, H. (2013). A biological motion toolbox for reading, displaying, and manipulating motion capture data in research settings. *Journal of Vision*, *13*(12), 7–7. https://doi.org/10.1167/13.12.7

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de-Wit, L., Wagemans, J., Eylen, L. Van, 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

Van de Cruys, S., Lemmens, L., Sapey-Triomphe, L., Chetverikov, A., Noens, I., & Wagemans, J. (2021). Structural and contextual priors affect visual search in children with and without autism. *Autism Research*, aur.2511. https://doi.org/10.1002/aur.2511

Van de Cruys, S., Perrykkad, K., Hohwy, J., Van, S., Cruys, D., Perrykkad, K., &

Hohwy, J. (2019). Explaining hyper-sensitivity and hypo-responsivity in autism with a common predictive coding-based mechanism. *Cognitive Neuroscience*, *10*(3), 164–166. https://doi.org/10.1080/17588928.2019.1594746

Van de Cruys, S., Van der Hallen, R., & Wagemans, J. (2017). Disentangling signal and noise in autism spectrum disorder. *Brain and Cognition*, *112*, 78–83. https://doi.org/10.1016/J.BANDC.2016.08.004

Van de Cruys, S., Vanmarcke, S., Van de Put, I., & Wagemans, J. (2018). The Use of Prior Knowledge for Perceptual Inference Is Preserved in ASD. *Clinical Psychological Science*, *6*(3), 382–393. https://doi.org/10.1177/2167702617740955

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods*, *45*(2), 576–594. https://doi.org/10.3758/s13428-012-0261-6

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2015). Meta-analysis of multiple outcomes: a multilevel approach. *Behavior Research Methods*, *47*(4), 1274–1294. https://doi.org/10.3758/s13428-014-0527-2

Van der Hallen, R., Evers, K., Brewaeys, K., Van den Noortgate, W., & Wagemans, J. (2015). Global processing takes time: A meta-analysis on local–global visual processing in ASD. *Psychological Bulletin*, *141*(3), 549–573. https://doi.org/10.1037/bul0000004

Van der Hallen, R., Manning, C., Evers, K., & Wagemans, J. (2019). Global Motion Perception in Autism Spectrum Disorder: A Meta-Analysis. *Journal of Autism and Developmental Disorders*, *49*, 4901–4918. https://doi.org/10.1007/s10803-019-04194-8

Van Eylen, L., Boets, B., Steyaert, J., Wagemans, J., & Noens, I. (2015). Local and Global Visual Processing in Autism Spectrum Disorders: Influence of Task

and Sample Characteristics and Relation to Symptom Severity. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-015-2526-2

Vetter, P., Edwards, G., & Muckli, L. (2012). Transfer of predictive signals across saccades. *Frontiers in Psychology*, *3*(JUN), 1–10. https://doi.org/10.3389/fpsyg.2012.00176

Vetter, P., Grosbras, M. H., & Muckli, L. (2015). TMS over V5 disrupts motion prediction. *Cerebral Cortex*, *25*(4), 1052–1059. https://doi.org/10.1093/cercor/bht297

Villalobos, M. E., Mizuno, A., Dahl, B. C., Kemmotsu, N., & Müller, R.-A. (2005). Reduced functional connectivity between V1 and inferior frontal cortex associated with visuomotor performance in autism. *NeuroImage*, *25*(3), 916–925. https://doi.org/10.1016/J.NEUROIMAGE.2004.12.022

Volkmar, F. R. (2011). Understanding the social brain in autism. *Developmental Psychobiology*, *53*(5), 428–434. https://doi.org/10.1002/dev.20556

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*, *371*, 2015373. https://doi.org/10.1098/rstb.2015.0373

Wadhera, T., & Kakkar, D. (2020). Multiplex temporal measures reflecting neural underpinnings of brain functional connectivity under cognitive load in Autism Spectrum Disorder Multiplex temporal measures reflecting neural underpinnings of brain functional connectivity under cognitive load in Autism Spectrum Disorder. *Neurological Research*, *42*(4), 327–337. https://doi.org/10.1080/01616412.2020.1726586

Walsh, K. S., McGovern, D. P., Clark, A., & O'Connell, R. G. (2020). Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Annals of the New York Academy of Sciences*, *1464*(1), 242–268. https://doi.org/10.1111/nyas.14321

Wang, L.-H., Chien, S. H. L., Hu, S.-F., Chen, T.-Y., Chen, H.-S., Hui, S., Chien, -Lin, Hu, S.-F. F., Chen, T.-Y. Y., & Chen, H.-S. S. (2015). Children with autism spectrum disorders are less proficient in action identification and lacking a preference for upright point-light biological motion displays. *Research in Autism Spectrum Disorders*, *11*(91), 63–76. https://doi.org/10.1016/j.rasd.2014.12.004

Wang, Y., Wang, L., Xu, Q., Liu, D., Chen, L., Troje, N. F., He, S., & Jiang, Y. (2018). Heritable aspects of biological motion perception and its covariation with autistic traits. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(8), 1937–1942. https://doi.org/https://dx.doi.org/10.1073/pnas.1714655115

Webster, M. A. (2011). Adaptation and visual coding. *Journal of Vision*, *11*(5), 3–3. https://doi.org/10.1167/11.5.3

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. Psychological Corporation.

Westfall, J. (2016). *PANGEA: Power ANalysis for GEneral Anova designs*. http://jakewestfall.org/publications/pangea.pdf

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wigham, S., Rodgers, J., South, M., McConachie, H., & Freeston, M. (2015). The interplay between sensory processing abnormalities, intolerance of uncertainty, anxiety and restricted and repetitive behaviours in autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(4), 943–952. https://doi.org/10.1007/s10803-014-2248-x

Williams, J. H. G., Waiter, G. D., Gilchrist, A., Perrett, D. I., Murray, A. D., & Whiten, A. (2006). Neural mechanisms of imitation and “mirror neuron”

functioning in autistic spectrum disorder. *Neuropsychologia*, *44*, 610–621. https://doi.org/10.1016/j.neuropsychologia.2005.06.010

World Health Organisation. (2020). *International Classification of Diseases, 11th Revision (ICD-11)* (11th ed.). https://icd.who.int/en

Yang, Y. J. D., Allen, T., Abdullahi, S. M., Pelphrey, K. A., Volkmar, F. R., & Chapman, S. B. (2017). Brain responses to biological motion predict treatment outcome in young adults with autism receiving Virtual Reality Social Cognition Training: Preliminary findings. *Behaviour Research and Therapy*, *93*, 55–66. https://doi.org/https://dx.doi.org/10.1016/j.brat.2017.03.014

Yang, Y. J. D., Sukhodolsky, D. G., Lei, J., Dayan, E., Pelphrey, K. A., & Ventola, P. (2017). Distinct neural bases of disruptive behavior and autism symptom severity in boys with autism spectrum disorder. *Journal of Neurodevelopmental Disorders*, *9*(1). https://doi.org/10.1186/s11689-017-9183-z

Yantis, S., & Nakama, T. (1998). Visual interactions in the path of apparent motion. *Nature Neuroscience*, *1*(6), 508–512. https://doi.org/10.1038/2226

Yon, D., de Lange, F. P., & Press, C. (2019). The Predictive Brain as a Stubborn Scientist. *Trends in Cognitive Sciences*, *23*(1), 6–8. https://doi.org/10.1016/J.TICS.2018.10.003

Yon, D., Zainzinger, V., de Lange, F. P., Eimer, M., & Press, C. (2020). Action biases perceptual decisions toward expected outcomes. *Journal of Experimental Psychology: General*, *undefined*(undefined), undefined. https://doi.org/10.1037/xge0000826