Check for updates

REVIEW ARTICLE

Prediction in Autism Spectrum Disorder: A Systematic Review of Empirical Evidence

Jonathan Cannon, Amanda M. O'Brien, Lindsay Bungert, and Pawan Sinha

According to a recent influential proposal, several phenotypic features of autism spectrum disorder (ASD) may be accounted for by differences in predictive skills between individuals with ASD and neurotypical individuals. In this systematic review, we describe results from 47 studies that have empirically tested this hypothesis. We assess the results based on two observable aspects of prediction: learning a pairing between an antecedent and a consequence and responding to an antecedent in a predictive manner. Taken together, these studies suggest distinct differences in both predictive learning and predictive response. Studies documenting differences in learning predictive pairings indicate challenges in detecting such relationships especially when predictive features of an antecedent have low salience or consistency, and studies showing differences in habituation and perceptual adaptation suggest low-level predictive processing differences in ASD. These challenges may account for the observed differences in the influence of predictive priors, in spontaneous predictive movement or gaze, and in social prediction. An important goal for future research will be to better define and constrain the broad domain-general hypothesis by testing multiple types of prediction within the same individuals. Additional promising avenues include studying prediction within naturalistic contexts and assessing the effect of prediction-based intervention on supporting functional outcomes for individuals with ASD. ***Autism Res*** *2021, 14: 604–630.* © 2021 International Society for Autism Research, Wiley Periodicals LLC.

Lay Summary: Researchers have suggested that many features of autism spectrum disorder (ASD) may be explained by differences in the prediction skills of people with ASD. We review results from 47 studies. These studies suggest that ASD may be associated with differences in the learning of predictive pairings (e.g., learning cause and effect) and in low-level predictive processing in the brain (e.g., processing repeated sounds). These findings lay the groundwork for research that can improve our understanding of ASD and inform interventions.

Keywords: autism spectrum disorder; prediction; learning; brain; perception

Introduction

Suppose you believe that your bus will come in 10 min. There can be several consequences of harboring such a prediction. You might modify your goal-directed behavior, increasing the pace of your walk if you are far from the bus-stop. If you are already at the stop and know that your bus is coming soon, your expectation might affect your perception: you might accidentally step onto the wrong bus, misperceiving it as your own. Your prediction might lead you to pay more attention to approaching vehicles around the bus's expected arrival time. If the bus does not arrive on time, your expectation might lead to an emotional response (understandable frustration) and rapid learning (never trust the CT2 bus to stick to its schedule!).

Thus, behavior, perception, emotion, attention, and learning can all be deeply impacted by the brain's systems for making predictions. By extension, differences between individuals across these dimensions could be manifestations of underlying differences in predictive processes. Indeed, such differences in predictive skills have recently been hypothesized to underlie autism spectrum disorder (ASD). In 2014, several independent groups of researchers proposed variations on a hypothesis of a broad, domain-general predictive impairment[1] in autism (PIA) [Lawson, Rees, & Friston, 2014; Sinha et al., 2014; Van de Cruys et al., 2014]. These proposals all attempted to explain a constellation of seemingly unrelated aspects

[1]The original theory uses the term "impairment." However, we would like to acknowledge that not all cognitive diversity associated with ASD necessarily results in functional impairment. We will use the term "impairment" sparingly throughout this manuscript and only to refer to situations in which we would expect the observed difference to cause difficulties for the individual. In doing so, we hope to help move the field towards language that actively works to destigmatize the ASD label.

From the Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA (J.C., A.M.O., L.B., P.S.); Program in Speech and Hearing Bioscience and Technology, Harvard University, Cambridge, Massachusetts, USA (A.M.O.)
Jonathan Cannon and Amanda M. O'Brien should be considered joint first author.
Received October 27, 2020; accepted for publication January 21, 2021
Address for correspondence and reprints: Amanda M. O'Brien, Program in Speech and Hearing Bioscience and Technology, Harvard University, 260 Longwood Avenue, Boston, MA 02115. E-mail: amanda_obrien@g.harvard.edu
Published online 11 February 2021 in Wiley Online Library (wileyonlinelibrary.com)
DOI: 10.1002/aur.2482
© 2021 International Society for Autism Research, Wiley Periodicals LLC.

of the ASD phenotype, including insistence on sameness, social interaction, sensory hypersensitivities, difficulties interacting with moving objects, and difficulties with theory of mind, in terms of underlying differences in predictive processes.

A vulnerability of any hypothesis that attempts to unify diverse aspects of ASD in cognitive terms lies in its tendency toward reductionism: the ASD diagnosis is shaped by the pragmatic needs of individuals spanning a wide range of genotypes and phenotypes, their families, and the medical professionals who serve them. Attempts to "explain" autism symptoms with a single unifying theory risk diminishing this diversity and complexity. If the PIA, which claims to account for only a subset of traits of ASD, can indeed navigate these perils of reduction, it offers significant promise. If it is backed by sufficient empirical support, this hypothesis could provide a deeper mechanistic account of the observed social and nonsocial symptoms of ASD, and thus contribute to the development of better therapies, accommodations, and even refined diagnostic criteria and tools [Constant, Bervoets, Hens, & Van de Cruys, 2020; Haker, Schneebeli, & Stephan, 2016]. It also has the potential to link the study of ASD with research on the neuroscience of prediction, informing a new understanding of the neurobiological differences that give rise to ASD and motivating the exploration of new targets for neuroactive medicines [Haker et al., 2016]. Indeed, several formulations of the PIA hypothesis draw heavily on the framework of "predictive processing," which grounds a prediction-based account of perception, learning, and action in a growing set of neurobiological mechanisms [Bastos et al., 2012; Friston, 2018].

Following the PIA hypothesis, multiple lines of autism research, some initiated by the hypothesis and some already in progress, began to adopt the language of predictive differences in ASD. As a result, the conversation on prediction in ASD is now broad and diverse, making it simultaneously rich in scientific content, but also difficult to summarize. The aim of this systematic review is to provide conceptual structure for this growing body of knowledge and use it to characterize what has been learned about prediction in ASD over the past six years and identify fruitful directions for further research.

What is Prediction?

The colloquial usage of the term 'prediction' encompasses a broad range of phenomena and behaviors that may have very different underlying brain mechanisms. In order to frame the discussion of hypothesized predictive impairments in ASD, and delineate the scope of our literature review, it is helpful to begin by attempting to define "prediction." For the purposes of this review, we consider a *predictive process* to be a behavioral or neural process that:

1. Occurs in response to an "antecedent" event or observation;
2. Is based on a learned association between that antecedent and a consequence (Fig. 1);
3. Directly affects the organism's neural or behavioral response upon the arrival of the consequence (or its absence).

The "association" between antecedent and consequence may include a specific time separation (Δt) and a degree of certainty associated with the probabilistic strength of the antecedent-consequent relationship. This association may be part of a collection of associations accumulated by the organism over time that together represent a probability distribution of possible consequences following the antecedent (or, more formally, a "conditional distribution" conditioned on the antecedent).

The requirement of an "antecedent" in our scope of prediction intentionally excludes predictions or expectations that are always present (i.e., "static" predictions, or general beliefs about the world). However, we opted to include a small collection of studies in our review that did not perfectly fit this definition due to having no clear antecedent. In these studies, the participants' neural or behavioral responses to stimuli over the course of the experiment were affected by "predictions" based entirely on the statistics of the preceding stimuli. We included these because they were relevant to several formulations of the PIA hypothesis, and because they were not far removed from an intuitive notion of "prediction."

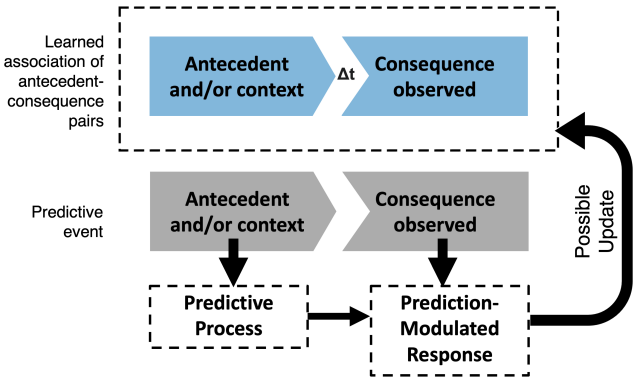

Figure 1. A simple visualization of the definition of prediction. The learned association (top) is built by aggregating observations of <antecedent, consequent> pairs. A predictive process (bottom) generates a prediction-modulated response upon being present with an antecedent, or a contextual setting. The congruence, or lack thereof, between the predicted and actual consequence supports a possible update to the antecedent-consequent association.

Our definition is limited to prediction about the future. This stands in contrast to certain definitions offered within the framework of predictive processing, for example, "estimates of unobserved or missing information on the basis of a model… (which) can be (but are not necessarily) future-oriented" [Teufel & Fletcher, 2020].[2] According to this definition, one might "predict" what is in an opaque box based on its weight and the sound it makes when it is shaken.

Since an organism's predictive faculties evolve through natural selection, the *prediction-modulated response* generally provides it some type of survival-relevant advantage[3]. A predictive process is like a bet placed on the predicted consequence: it can produce a win if the predicted consequence comes, but involves an expenditure of resources that may amount to a loss if it does not. The *prediction-modulated response* to the predicted consequence may take many forms, including a usefully biased percept, attenuated neural firing, or a speeded motor response. The response to an alternative, unpredicted consequence (a "prediction error") may include a counterproductively biased percept, heightened neural firing, a slowed motor response, and/or rapid learning.

Figure 1 presents a schematic definition of prediction. A simple example illustrates the use of a learned association to initiate a predictive process and the resultant prediction-modulated response. You may have a learned association of what it sounds like when your supervisor is walking down the hall to your office. This model includes a collection of associations between the sounds you are hearing and the consequence of your supervisor arriving at your office. These associations are based on accumulated experiences from your time working with your supervisor. You may also be aware of the length of time it will take for your supervisor to walk to your office. Based on these learned associations, hearing the antecedents within the context of your office will likely initiate a predictive process about the arrival of the supervisor to your office. This predictive process may be evidenced by your behaviors (e.g., closing internet windows that are unrelated to work, and picking up a document for review). When you see your supervisor standing at your office door, you may demonstrate a prediction-modulated response (e.g., having a faster reaction time when saying hello). If your prediction is incorrect, and it is not your supervisor but rather your colleague, you will use this information to update your antecedent-consequent (e.g., your colleague and supervisor approach in similar sounding ways), and the probabilistic strength of the antecedent-consequent relationship will weaken.

Including a consideration of "benefit to organism" in the definition of prediction helps distinguish between processes that merely affect the organism's response to a consequence in terms of the organism "expecting" the consequence—without the prospect of some benefit from those that confer an adaptive advantage. A focus on the benefits of predictive processes also makes clear the pragmatic value of an investigation into prediction in autism: predictive processes are only "impaired" in ASD insofar as they deprive the individual with ASD of desired benefits of prediction, and their improvement through therapy or other means is to be measured in terms of accessing these benefits.

In order to organize the wealth of research related to prediction in ASD, we have identified three separate types of observable *evidence* that prediction is occurring:

Behavioral evidence (e.g., moving one's eyes to the door when one hears approaching footsteps).

Perceptual report or judgment (e.g., identifying a cell phone as a gun when it is pulled out by a stranger in a dark alleyway).

Neural evidence (e.g., minimizing the brain's electrophysiological response to an expected word at the end of a sentence).

Introspection-driven overt response (e.g., having a person describe what chess move their opponent will make next).

These processes further fragment according to the nature of the experiences that give rise to the pairing of the antecedent and consequence (i.e., the *basis* of the predictive model), which may range from a single observation to a lifetime worth of experience, and may be driven by explicitly conveyed information or inferences from experience.

We review the past six years of research on prediction in ASD from two vantage points. We first group the research by the basis of the predictive model (i.e., how the pairing of antecedent and consequence is established). This allows us to investigate whether the process of learning predictive models is specifically impaired in ASD. We then group the research by the type of predictive process under consideration, separating the studies based on whether they observe the participant predictively optimizing their behavior, perceptual

[2]However, note that even "predictions" about hidden aspects of the present or past can only be tested and evaluated on the basis of observations in the future, and thus can be reframed as future-oriented predictions about the outcomes of these observations.

[3]The notion of "prediction" as securing some value or advantage for the organism is similar to the framing of prediction in Schubotz [2016] in terms of a "manifestation of an energetic investment" reflecting the exploitation of predictive information and a "manifestation of a benefit" due to the predicted event not having to be awaited or fully processed. It is also in line with the predictive processing framework, which is grounded in a "free-energy principle" that frames all prediction (and all action) as serving to minimize a long-term average of "surprise," defined as the deviation of the organism's state from a narrow preferred distribution [Friston, 2010]. Minimizing surprise involves securing basic needs for the organism (e.g., food) and also gathering information about the world ("epistemic value"). We acknowledge, however, that predictive processes do not necessarily provide an advantage. For example, someone may experience discomfort and anxiety upon hearing the voice of someone who has caused them pain in the past.

response, neural state, or simply document reports of their predictions. This affords us the opportunity to assess whether any of these types of prediction are specifically affected in ASD.

Methods

Our systematic review followed Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [Moher et al., 2010]. We began with a literature search of the PubMed database using the following keywords within the title or abstract of articles published between Jan 1, 2014 and June 30, 2020: ("Autism" OR "Autism spectrum" OR "Asperger's" OR "PDD-NOS") AND ("Prediction" OR "Expectation"). We also included all articles that cited the Sinha et al. [2014] and/or the Van de Cruys et al. [2014] theoretical papers, using the *Publish or Perish* software program. We removed all duplicates from the search and excluded any papers that were not written in English, which left 711 papers for eligibility screening (Fig. 2).

We used the following inclusion criteria: (a) articles were peer-reviewed and included an experimental design, thus excluding purely theoretical and computational papers; (b) articles included participants with a diagnosis of ASD, thus excluding studies that had only neurotypical individuals who were rated on ASD symptomatology (e.g., via the Autism Quotient); and (c) papers included a methodology that provided a measurement of prediction, as defined above. All articles and records were managed privately through the Open Science Foundation website.

We double-coded 15% of the papers to ensure consistency. We discussed discrepancies until agreement reached 100%. Admittedly, we may have missed empirical studies exploring phenomena that are closely related to predictive processes, but do not use the terminology of prediction that we used in our search. Notwithstanding this limitation, we are encouraged that our search generated a large corpus of papers that could form the basis of a substantive review.

The 47 remaining articles were categorized according to both (a) the *basis* of the predictive model (i.e., the genesis of the pairing between the antecedent and consequence); and (b) the behavioral, perceptual, or neural *evidence* measured as an indicator of a predictive process. Each study is discussed below in the context of its basis (in Section 3.1) and then discussed again in the context of its evidence (in Section 3.2). Table 1 lists the basis and evidence for each of the included articles.

A note on terminology: Although we abide by the referential conventions of the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5), we acknowledge that the term "autistic person" is preferred by many people on the spectrum.

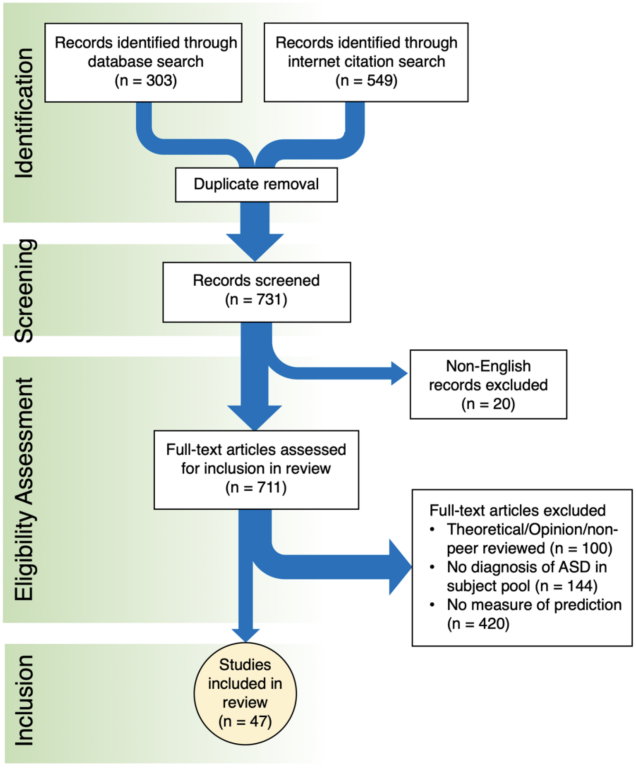

Figure 2. Flow diagram of article selection following the template outlined by PRISMA guidelines [Moher et al., 2010].

Results

Results Organized by Bases for Prediction

All included studies were categorized by basis of the predictive model (i.e., what information is the prediction based on?) used by participants in the study. Seven bases for prediction were identified.

Stimulus constancy. The simplest possible antecedent/consequence pairing is the expectation that a stimulus will be followed by a similar stimulus. Thirteen studies looked at predictions based on the expectation of stimulus constancy. Of these, eight looked for evidence of prediction in the form of habituation (i.e., the attenuation of neural or behavioral response over repetitions of the same stimulus). The predictability of a stimulus sequence is directly proportional to the observed neural habituation [Herry et al., 2007; Sokolov, 1960; McDaniel, 1964; Turk-Browne et al., 2008]. This habituation may be related to the conservation of neural resources and be important for optimizing one's response to unexpected input. Sinha et al. [2014] described the potential consequences of reduced habituation for individuals with ASD, which include hypersensitivity to sensory input and

Table 1. All Included Articles, Study Demographics, Stimuli Type, Evidence and Basis for Prediction, and Primary Study Findings

Author	Title	ASD	NT	Modality	Evidence of prediction (in paper)	Bases for prediction	Social aspects	Findings
Amoruso et al., 2019	Contextual Priors Do Not Modulate Action Prediction in Children with Autism	*n* = 24 Age = 8.66 (1.63) IQ = 107.91 (10.62) Diagnosis based on medical review.	*n* = 24 Age = 9.04 (1.08) IQ = 102.08 (10.2)	Visual	Perceptual prior	Associations learned over trials	Stimuli: videos of social and asocial human action	ASD group did not use statistical associations of context and action to predict what action was being performed when movement information was obscured, but NT group did.
Balsters et al., 2017	Disrupted Prediction Errors Index Social Deficits in Autism Spectrum Disorder	*n* = 16 Age = 20.97 (3.17) IQ = 115.5 (10.61)	*n* = 20 Age = 22.17 (5.2) IQ = 118 (11)	Visual	Reported surprise, fMRI: Prediction-related evoked activity	Instructed rules	Perspective taking	ASD group demonstrates reduced awareness of prediction accompanied by reduced prediction-error-related frontal BOLD activity when predicted reward is experienced by others.
Barzy et al., 2019	Autistic Adults Anticipate and Integrate Meaning Based on The Speaker's Voice: Evidence from Eye-Tracking and Event-Related Potentials	*n* = 25 Age = 34.4 (10.78) IQ = 106.88 (15.14)	*n* = 24 Age = 33.04 (16.88) IQ = 101.79 (10.91)	Visual	Spontaneous gaze, EEG: Prediction-related evoked activity (N400)	Social experience	Stimuli: voices	When a specific target is expected based on speaker identity, ASD group looked to target later but displayed consistent N400 to expectation violation.
Braukmann et al., 2018	Action Prediction In 10-Month-Old Infants at High and Low Familial Risk for Autism Spectrum Disorder	*High Risk Infants* *n* = 29 Age = 10.18 mos (0.51) MSEL = 92.66 (13.93)	*Low Risk Infants* *n* = 23 Age = 10.13 mos (0.41) MSEL = 96.57 (13.33)	Visual	Spontaneous gaze	Associations learned over lifetime	Stimuli: recognizable actions	Infants at risk of ASD showed equal gaze prediction of end point of recognizable action and equal learning of end point over trials
Chambon et al., 2017	Reduced Sensitivity to Social Priors During Action Prediction in Adults with Autism Spectrum Disorders	*n* = 18 Age = 35.7 (7.7) IQ = 104.9 (18.6)	*n* = 20 Age = 34.8 (6.4) IQ = 107.7 (7.9)	Visual	Perceptual prior	Social experience	Stimuli: videos of social and nonsocial actions	ASD group displayed differences in predictions based on social priors. Predictions based statistical regularities over the course of the experiment, and the process of learning these regularities, were not significantly different.
Deschrijver et al., 2016	Behavioral Measures of Implicit Theory of Mind in Adults with High Functioning Autism	*n* = 13 Age = 32.95 (6.26) IQ = 110.95 (14.64)	*n* = 12 Age = 31.89 (6.26) IQ = 117.89 (14.23)	Visual	Response time	Object permanence	Perspective taking	ASD group RTs were more influenced by the expectation of object permanence. There was a weak trend toward RTs less influenced by expectations of an actor on the scene.
Ego et al., 2016	Behavioral Characterization of Prediction and Internal Models in Adolescents with Autistic Spectrum Disorders	*n* = 40 Age = 16.4 (4.2) IQ = 98 (36)	*n* = 35 Age = 15.7 (4.1) IQ = 108 (38)	Visual	Spontaneous gaze	Kinematic/ temporal extrapolation		ASD group was equally able to visually track a moving object through an occlusion.

Ewbank et al., 2017	Repetition Suppression and Memory for Faces Is Reduced in Adults with Autism Spectrum Conditions	*n* = 15 Age = 31.8 (9.2) IQ = 126 (11.8) Diagnosis based on DSM-IV.	*n* = 15 Age = 28.1 (7.6) IQ = 128.2 (10.5)	Visual	fMRI: Prediction-related evoked activity	Stimulus constancy	Stimuli: social (faces) and nonsocial (shapes)	ASD group demonstrated reduced repetition suppression for faces, but not shapes.
Font-Alaminos et al., 2020	Increased Subcortical Neural Responses to Repeating Auditory Stimulation in Children with Autism Spectrum Disorder	*n* = 17 Age = 9.1 (1.7) IQ = 103.8 (20.3)	*n* = 18 Age = 8.8 (1.9) IQ = 111.4 (13.9)	Auditory	EEG: Prediction-related evoked activity (frequency-following)	Stimulus constancy		ASD group showed increased frequency-following subcortical response to repeated sounds, suggesting reduced attenuation by prediction.
Fulceri et al., 2018	Interpersonal Motor Coordination During Joint Actions in Children with And Without Autism Spectrum Disorder: The Role of Motor Information.	*n* = 11 Age = 7.82 (1.32) IQ = 100.45 (11.76)	*n* = 11 Age = 7.57 (0.71) IQ = —	Visual	Coordinated movement	Kinematic/ temporal extrapolation	Dyadic interaction task	ASD group showed impaired joint action when they had to rely on partner's kinematic cues in the absence of goal cues.
Ganglmayer et al., 2020	Do Children and Adults with Autism Spectrum Condition Anticipate Others' Actions as Goal-Directed? A Predictive Coding Perspective	*Children* *n* = 30 Age = 9.73 (1.86) IQ = 109.93 (14.83) Diagnosis based on medical review.	*Children* *n* = 29 Age = 9.34 (1.72) IQ = 105.59 (13.88)	Visual	Spontaneous gaze	Statistics accumulated over trials	Inference of agent's goal	The ASD group eventually learned to use saccades to anticipate agent's goal but needed more learning trials to infer a goal.
Ganglmayer et al., 2020		*Adolescents* *n* = 19 Age = 15.05 (1.54) IQ = 103.32 (22.12)	*Adolescents* *n* = 19 Age = 15.11 (1.33) IQ = 115.53 (17.32)					
Ganglmayer et al., 2020		*Adults* *n* = 22 Age = 33.86 (13.10) IQ = 103.14 (21.65)	Adults *n* = 24 Age = 38.46 (14.55) IQ = 108.25 (14.88)					
Gonzalez-Gadea et al., 2015	Predictive Coding in Autism Spectrum Disorder and Attention Deficit Hyperactivity Disorder	*n* = 24 Age = 10.38 (1.97) fluid intelligence on Raven's progressive matrices test = 39.63 (9.83) Diagnosis based on DSM-V.	*n* = 19 Age = 11.63 (2.43) Fluid intelligence on Raven's progressive matrices test = 40.16 (8.20)	Auditory	EEG: Prediction-related evoked activity (MMN, P3)	Stimulus constancy		ASD group showed a reduced P3 response to task-irrelevant deviants but an increased response to task-relevant deviants. MMN responses to deviants were identical.
Goris et al., 2018	Sensory Prediction Errors Are Less Modulated by Global Context in Autism Spectrum Disorder	*n* = 18 Age = Adults IQ > 80	*n* = 24 Age = Adults IQ > 80	Auditory	EEG: Prediction-related evoked activity (MMN)	Statistics accumulated over trials		MMNs of ASD group were less modulated by likelihood of deviant.

(Continues)

Table 1. Continued

Author	Title	ASD	NT	Modality	Evidence of prediction (in paper)	Bases for prediction	Social aspects	Findings
Greene et al., 2019	Social and Nonsocial Visual Prediction Errors in Autism Spectrum Disorder	*n* = 25 Age = 14.78 (1.62) IQ = 104.40 (18.22)	*n* = 18 Age = 14.81 (2.08) IQ = 109.28 (14.88)	Visual	Spontaneous gaze	Instructed statistical regularities	Stimuli: social (faces) and nonsocial (objects)	Instructed to look at stimulus after appearance and informed of cues that statistically predicted position, ASD group showed fewer position-predictive saccades.
Grisoni et al., 2019	Prediction and Mismatch Negativity Responses Reflect Impairments in Action Semantic Processing in Adults with Autism Spectrum Disorders	*n* = 20 Age = 38 (10.3) IQ = 119.5 (8.4) Diagnosis based on DSM-IV.	*n* = 22 Age = 31.9 (11.1) IQ = 116.8 (9.5)	Auditory	EEG: Prediction-related evoked activity (MMN), EEG: Predictive activity	Associations learned over lifetime (word-sound)	Stimuli: action and non-action words and sounds	ASD group showed reducted predictive potential for action sounds, reduced MMN for action standards, and reduced MMN modulation by semantic congruency for action standards.
Karaminis et al., 2016	Central Tendency Effects in Time Interval Reproduction in Autism	*n* = 23 Age = 12.33 (1.83) IQ = 100.30 (15.72)	*n* = 23 Age = 11.67 (1.58) IQ = 102.78 (11.26)	Visual	Perceptual prior	Statistics accumulated over trials		ASD group showed impaired temporal discrimination, and relatively weaker influence of learned priors on duration given their poorer discrimination
Karaminis et al., 2020	Adaptation to The Speed of Biological Motion in Autism	*n* = 19 Age = 14.15 (2.84) IQ: 104.32 (16.57)	*n* = 19 Age = 13.93 (3.80) IQ: 105.58 (12.65)	Visual	Perceptual adaptation	Stimulus constancy		ASD group showed no differences in adaptation to biological and nonbiological movement after controlling for visual attention.
Kinard et al., 2020	Neural Mechanisms of Social and Nonsocial Reward Prediction Errors in Adolescents with Autism Spectrum Disorder	*n* = 22 Age = 14.66 (1.65) IQ = 100.95 (17.04)	*n* = 20 Age = 14.75 (2.07) IQ = 51.83 (10.50)	Visual	fMRI: Prediction-related evoked activity	Instructed statistical regularities	Reward stimuli: faces and pleasing objects	In the social reward condition, ASD group showed decreased activation in the right precentral gyrus and increased activation in the right frontal pole with positive (vs. negative) reward prediction errors, and greater activation in the right anterior cingulate gyrus and bilateral precentral gyrus with larger magnitude errors. No differences were observed in the non-social condition.
Kolesnik et al., 2019	Increased Cortical Reactivity to Repeated Tones At 8 Months in Infants with Later ASD	*High Risk—ASD* *n* = 14 Age (months) = 8.83 (0.82 m); 38.56 (1.71) MSEL Receptive: 47.4 (8.9); 39.6 (16.5)	*Low Risk Infants* *n* = 14 Age (months) = 9.29 (0.84) MSEL Receptive: 49 (11.05); 59.08 (9.5)	Auditory	EEG: Prediction-related evoked activity (oscillations)	Stimulus constancy		Infants later diagnosed with ASD showed reduced repetition suppression and increased phase locking to auditory stimuli.
Kolesnik et al., 2019		*High Risk—Atypical* *n* = 21 Age (months) = 8.90 (0.69); 38.69 (1.87) MSEL Receptive = 49.11 (7.3); 44.11 (14.78)	High Risk—NT *n* = 44 Age (months) = 9.10 (0.84); 38.84 (1.60) MSEL Receptive = 50.98 (9.65); 56.91 (9.71)					

Krogh-Jespersen et al., 2018	Goal Prediction In 2-Year-Old Children with And Without Autism Spectrum Disorder: An Eye-Tracking Study.	*n* = 36 Age (months) = 27;27 MSEL = 64.9 (12.6)	*n* = 19 Age (months) = 26;23 MSEL = 102.2 (15.4)	Visual	Spontaneous gaze	Association learned by single exposure	Inference of agent goal	ASD group did not differ from NT group in performance of anticipatory saccades based on kinematic cues. In absence of kinematic cues, ASD group made anticipatory saccades to previous target location, whereas age-matched NT group made saccades based on inferred agent goal.
Król & Król, 2019	The World as We Know It and The World as It Is: Eye-Movement Patterns Reveal Decreased Use of Prior Knowledge in Individuals with Autism	*n* = 21 Age = 16.00 (4.92) IQ = 109.81 (17.65)	*n* = 23 Age = 16.11 (2.66) IQ = 112.65 (10.94)	Visual	Spontaneous gaze	Familiarity with complex stimulus	Stimuli: social and asocial images	On Mooney images, ASD group showed less optimization of eye scan pattern over learning and less scan path stability across presentations. There was no significant group difference between social and asocial images.
Kunchulia et al., 2017	Time-Based Event Expectancies in Children with Autism Spectrum Disorder.	*n* = 9 Age = 8.15 (1.7) IQ = 85.2 (14.1)	*n* = 10 Age = 8.9 (1.95) IQ = 104.5 (6.47)	Visual	Response time	Associations learned over trials		ASD group demonstrated use of temporal regularities to make predictions more similar to NT adults than age-matched NT controls.
Kunchulia et al., 2020	Children with Autism Spectrum Disorder Show Increased Sensitivity to Time-Based Predictability	*Task 1* *n* = 11 Age = 9 (1.9) IQ = 86.4 (12)	*Task 1* *n* = 11 Age = 8.5 (1.8) IQ = 103.2 (9.8)	Visual	Response time	Associations learned over trials		Children with ASD made more effective use of predictive value of pre-target interval duration for speeding reaction time.
Kunchulia et al., 2020		*Task 2* *n* = 11 Age = 9 (1.9) IQ = 86.4 (12)	*Task 2* *n* = 11 Age = 9 (1.9) IQ = 86.4 (12)					
Landa et al., 2016	Ready, Set, Go! Low Anticipatory Response During A Dyadic Task in Infants at High Familial Risk for Autism	*High Risk Infants (6 mos)* *n* = 66 Age (mos) = 6.7 (0.56) MSEL VR = 6.7 (0.97)	*Low Risk Infants (6 mos)* *n* = 43 Age (mos) = 6.7 (0.51) MSEL VR = 7.0 (1.0)	Visual	Spontaneous movement	Kinematic/ temporal extrapolation	Dyadic interaction task	Infants at high risk of ASD were less likely to demonstrate anticipatory motor response to the approaching ball than controls. The difference in motor response was not due to differences in visual attention.
Lawson et al., 2017	Adults with Autism Overestimate the Volatility of The Sensory Environment	*n* = 27 Age = 35.5 IQ = 117	*n* = 26 Age = 36 IQ = 120	Visual	Response time	Associations learned over trials		In task with changing contingencies, the ASD group demonstrated decreased behavioral and pupillometric indicies of surprise for unlikely stimuli. Computational modeling indicated that the ASD group relearned contingencies at a rate that was inflexible to volatility of the contingencies, and that pupillary response was more, not less, sensitive to suprise relative to learned expectations.

(Continues)

Table 1. Continued

Author	Title	ASD	NT	Modality	Evidence of prediction (in paper)	Bases for prediction	Social aspects	Findings
Lieder et al., 2019	Perceptual Bias Reveals Slow-Updating in Autism and Fast-Forgetting in Dyslexia	*n* = 37 Age = 26.1 (4.94) IQ = – Diagnosis based on DSM-V.	*n* = 33 Age = 24.1 (1.9) IQ = –	Auditory	Perceptual prior	Statistics accumulated over trials		ASD group showed reduced perceptual bias from previous stimuli, with especially weak bias from most recent stimuli.
Manning et al., 2017	Children on The Autism Spectrum Update Their Behaviour In Response to A Volatile Environment	*n* = 24 Age = 9.92 (2.0) IQ = 102.94 (15.46)	*n* = 32 Age = 9.17 (1.83) IQ = 107.28 (10.66)	Visual	Reward-oriented decision	Associations learned over trials		ASD group did not demonstrate a group-level difference in performance or in adapting learning rate to volatility.
Millin et al., 2018	Reduced Auditory Cortical Adaptation in Autism Spectrum Disorder	*n* = 21 Age = 23 IQ = 114	*n* = 27 Age = 24 IQ = 112	Auditory and Visual	fMRI: Prediction-related evoked activity	Stimulus constancy		ASD group showed reduced adaptation to repeated stimuli, but moreso to auditory than to visual stimuli and moreso to when interstimulus interval was fixed than when it was random.
Mosner et al., 2019	Neural Mechanisms of Reward Prediction Error in Autism Spectrum Disorder	*n* = 16 Age = 19.50 (2.07) IQ = 106.93 (18.22)	*n* = 14 Age = 31.47 (9.99) IQ = 110.33 (14.16)	Visual	fMRI: Prediction-related evoked activity	Associations learned over trials		ASD group showed greater activation in the left paracingulate gyrus with positive (vs. negative) reward prediction errors, and greater activation in the left insula and right frontal pole with larger magnitude reward prediction errors.
Northrup et al., 2017	Response to Changing Contingencies in Infants at High and Low Risk for Autism Spectrum Disorder	*High Risk Infants* *n* = 39 Age = Infants IQ = –	*Low Risk Infants* *n* = 17 Age = Infants IQ = –	Auditory and Visual	Spontaneous movement	Associations learned in single exposure		At 10 months, infants at high risk for ASD shook silent rattle less than controls, indicating reduced expectation based on recent rattle experience. No difference between groups at 6 months.
Palumbo et al., 2015	Atypical Emotional Anticipation in High-Functioning Autism.	*Exp. 1*: *n* = 16 Age = 20.3(3.2) IQ = 117.1(10.9)	Exp. 1: *n* = 18 Age = 20.1(4) IQ = 113.4(8.4)	Visual	Perceptual judgment	Social experience	Stimuli: facial expressions	ASD group was equally susceptible to "overshoot" effects in sequences of emotional faces, but further experiments showed that this effect was driven by basic stimulus contrast rather than social priors.
Palumbo et al., 2015		*Exp. 2*: *n = 17* *Age = 22.1 (7.2)* *IQ = 115.9 (8.1)*	Exp. 2: *n* = 17 Age = 22.1 (5.5) IQ = 114.5 (7.7)	Visual				
Ruiz-Martínez et al., 2020	Impaired P1 Habituation and Mismatch Negativity in Children with Autism Spectrum Disorder	*n* = 16 Age = 8.96 ± 1.01 SD Kaufman Brief Intelligence Test = 50.26 (38.99)	*n* = 15 Age = 8.86 ± 1.77 SD Kaufman Brief Intelligence Test = 85 (15.99)	Auditory	EEG: Prediction-related evoked activity (P1, MMN)	Stimulus constancy	Stimuli: human and non-human sounds	ASD group showed reduced P1 habituation and decreased MMN.

Schuwerk et al., 2016	Cognitive Mechanisms Underlying Action Prediction in Children and Adults with Autism Spectrum Condition	*Children* *n* = 26 Age = 9.7 (1.9) IQ = 105.4 (16.1)	*Children* *n* = 24 Age = 9.7 (1.9) IQ = 110.1 (15.9)	Visual	Spontaneous gaze	Statistics accumulated over trials	Inference of agent's goal	ASD group showed fewer action-anticipatory saccades. Anticipatory gaze in NT groups was based on statistical learning, indicating reduced use of statistical information to inform anticipatory gaze in ASD group.
Schuwerk et al., 2016		*Adults* *n* = 18 Age = 37.1 (11.5) IQ = 108.1 (21.4)	*Adults* *n* = 17 Age = 35.9 (14.1) IQ = 109.1 (16.1)					
Sheppard et al., 2016	Difficulties Predicting Time-To-Arrival in Individuals with Autism Spectrum Disorders	*n* = 23 Age = 18.55 (1.79) IQ = 99 (12.15) Diagnosis based on DSM-V.	*n* = 21 Age = 18.83 (2.25) IQ = 99.48 (15.48)	Visual	Reported prediction	Kinematic/ temporal extrapolation		ASD group demonstrated impaired judgements of arrival time, but only for straight roads with limited visual self-motion cues.
Tam et al., 2017	Altered Behavioral and Amygdala Habituation in High-Functioning Adults with Autism Spectrum Disorder: An Fmri Study	*n* = 22 Age = 34.1 (11.5) IQ = 106.7 (14.0)	*n* = 24 Age = 36.2 (11.0) IQ = 113.9 (11.0)	Visual	Response time, fMRI: Prediction-related evoked activity	Stimulus constancy	Distractor stimuli: faces	ASD group demonstrated deficits in neural and response time habituation.
Tewolde et al., 2018	Visual Motion Prediction and Verbal False Memory Performance in Autistic Children	*n* = 30 Age = 11.16 (2.23) IQ = 105.67 (15.32)	*n* = 30 Age = 10.48 (2.18) IQ = 11.40 (14.74)	Visual	Coordinated movement	Kinematic/ temporal extrapolation		ASD group did not significantly differ from NT group in precision of extrapolating linear movement or accumulation.
Thillay et al., 2016	Atypical Brain Mechanisms of Prediction According to Uncertainty in Autism.	*n* = 12 Age = 21.33 (11) IQ = 101 (5)	*n* = 12 Age = 21.58 (11) IQ = None, took 2 matrix tests	Visual	Response time, EEG: Prediction-related evoked activity (CNV, mu)	Instructed statistical regularities		ASD group reaction times and EEG mu rhythm were less modulated by target predictability. Faster reaction times and larger CNVs to unpredictable targets indicated "overanticipation" of unpredictable targets.
Turi et al., 2015	Children with Autism Spectrum Disorder Show Reduced Adaptation to Number	*n* = 16 Age = 10.3 (2.1) IQ = 107.15 (15.31)	*n* = 18 Age = 11 (2.1) IQ = 115.44 (7.55)	Visual	Perceptual adaptation	Stimulus constancy		The ASD group showed less perceptual adaptation to numerosity.
Turi et al., 2016	No Rapid Audiovisual Recalibration in Adults on The Autism Spectrum	*n* = 16 Age = 29.2 (5.2) IQ = 112.0 (10.32)	*n* = 16 Age = 27.1 (2.83) IQ = 112.7 (11.44)	Audiovisual	Perceptual adaptation	Stimulus constancy		ASD group demonstrated reduced adaptation to audio-visual asynchrony.
Utzerath et al., 2018	Adolescents with Autism Show Typical Fmri Repetition Suppression, But Atypical Surprise Response	*n* = 22 Age = 12–18 IQ > 85	*n* = 22 Age = 12–18 IQ > 85	Visual	fMRI: Prediction-related evoked activity	Associations learned over trials, Stimulus constancy		ASD group displayed similar repetition suppression. However, the ASD group showed less visual cortex suppression for images expected based on repeated pairings.

(Continues)

Table 1. Continued

Author	Title	ASD	NT	Modality	Evidence of prediction (in paper)	Bases for prediction	Social aspects	Findings
van Boxtel et al., 2016	Intact Recognition, But Attenuated Adaptation, For Biological Motion in Youth with Autism Spectrum Disorder	*n* = 16 Age = 14.04 (2.24) IQ = 101.50 (20.14)	*n* = 17 Age = 13.32 (3.53) IQ = 112.24 (14.01)	Visual	Perceptual adaptation	Stimulus constancy		ASD group demonstrated reduced perceptual adaptation to motion stimuli.
Van de Cruys et al. 2018	The Use of Prior Knowledge for Perceptual Inference Is Preserved In ASD	*n* = 23 Age = 14.04 (1.49) IQ = 104.33 (8.74)	n = 24 Age = 14.38 (1.28) IQ = 105.96 (8.69)	Visual	Perceptual prior	Familiarity with complex stimulus		ASD group did not significantly differ in use of prior exposure to inform performance on Mooney image recognition task.
van Laarhoven et al., 2019	Electrophysiological Alterations in Motor-Auditory Predictive Coding in Autism Spectrum Disorder.	*n* = 30 Age = 18.55 (2.13) IQ = 103.00 (16.47)	*n* = 30 Age = 18.83 (1.32) IQ = 111.97 (11.49)	Auditory	EEG: Prediction-related evoked activity (N1)	Self-initiation		ASD group displayed N1 response to self-initiated tones while the NT group did not.
van Laarhoven et al., 2020	Atypical Visual–Auditory Predictive Coding in Autism Spectrum Disorder: Electrophysiological Evidence from Stimulus Omissions	*n* = 29 Age = 18.64 (2.11) IQ = 103.03 (16.67) Diagnosis based on medical review.	*n* = 29 Age = 18.93 (1.22) IQ = 112.07 (11.68)	Audiovisual	EEG: Prediction-related evoked activity (ON1)	Associations learned over lifetime	Stimuli: hand claps	ASD group demonstrated larger omission-evoked responses than the NT group when sound of a visible handclap was omitted.
Vivanti et al., 2018	Attention to Novelty Versus Repetition: Contrasting Habituation Profiles in Autism and Williams Syndrome	*n* = 39 Age in Months = 44.16 (11.56) Mullen Scales of Early Learning (MSEL) = 62.08 (26.79)	*n* = 19 Age in Months = 49.45 (11.49) Mullen Scales of Early Learning (MSEL) = 105.29 (14.31)	Visual	Spontaneous gaze	Stimulus constancy		ASD group demonstrated reduced attentional (gaze) habituation.
Von Der Lühe et al., 2016	Interpersonal Predictive Coding, Not Action Perception, Is Impaired in Autism	*n* = 16 Age = 41.56 (9.15) IQ = 116.88 (15.59) Diagnosis based on DSM-V.	*n* = 16 Age = 36.19 (12.11) IQ = 115.31 (8.43)	Visual	Perceptual prior	Social experience	Stimuli: degraded images of social and asocial human action	ASD group demonstrated reduced utilization of social cues to enhance visual discrimination.
Westerfield et al., 2015	Tracking the Sensory Environment: An ERP Study of Probability and Context Updating In ASD.	*n* = 16 Age = 32.3 (10) IQ = 103.6 (16)	*n* = 16 Age = 32.8 (10) IQ = —	Visual	EEG: Prediction-related evoked activity (P3)	Social aspects		ASD P3 responses to targets showed greater sensitivity to probability, but P3 responses to non-targets showed reduced sensitivity to probability.

Note. ASD group differences relative to an NT group unless otherwise stated. ASD diagnosis is based on the Autism Diagnostic Observation Schedule (ADOS) unless otherwise indicated.

stress associated with unceasingly salient repetitive stimuli.

Seven of the eight studies found some type of reduced habituation in the ASD group compared to the neurotypical (NT) group. This included decreased neural habituation to repeated tones [Font-Alaminos et al., 2020; Kolesnik et al., 2019; Millin et al., 2018; Ruiz-Martínez et al., 2020] and faces [Ewbank et al., 2017; Tam et al., 2017] but not to other images [Millin et al., 2018; Utzerath, Schmits, Buitelaar, & de Lange, 2018] as measured by EEG and fMRI. Overall, these results suggest widespread but stimulus-specific reduction in neural habituation to repetition. Tam et al. [2017] also found decreased habituation of reaction time (RT) over repetitions of a working memory task involving faces, while Vivanti et al. [2018] found decreased habituation of gaze fixation frequency to repeating shapes.

Four more studies [Karaminis, Arrighi, Forth, Burr, & Pellicano, 2020; Turi et al., 2015; Turi, Karaminis, Pellicano, & Burr, 2016; van Boxtel, Dapretto, & Lu, 2016] looked at perceptual adaptation to recent stimuli. In Section 3.2.2 (Perceptual Adaptation), we describe how this effect can be understood as an effect of predicting that subsequent stimuli will be similar.

The final study in this group, and one of the aforementioned studies, used a classic oddball paradigm, in which neural responses, typically EEG, are measured when the expectation of stimulus repetition is disrupted by deviants. Ruiz-Martínez et al. [2020] found reduced EEG response to deviants during passive listening. This result is put in a larger context by Gonzalez-Gadea et al. [2015], who found reduced EEG responses to task-irrelevant deviants, but increased responses to task-relevant deviants. Rather than indicating weaker expectations based on the prediction of stimulus constancy, these results highlight the importance of task instructions and task-relevance in the study of prediction in ASD.

Kinematic and temporal extrapolation. Five studies looked at predictions based primarily on the extrapolation of a temporal or kinematic process forward in time. Of these, two reported differences in anticipatory action [Fulceri et al., 2018; Landa, Haworth, & Nebel, 2016], two reported no impairment in anticipatory behavior (eye movements [Ego et al., 2016] or instructed timed response [Tewolde, Bishop, & Manning, 2018]), and one reported difference in predictive judgment that was limited to the case in which participants had reduced visual cues [Sheppard, van Loon, Underwood, & Ropar, 2016]. Notably, the two studies reporting more significant differences both measured anticipatory motor coordination without specific instructed goals, indicating that these results may be due to a reduced inclination to act anticipatorily rather than a fundamental impairment in kinematic or temporal extrapolation.

Two more studies looked at predictions based on kinematic/temporal extrapolation in combination with other information [Amoruso et al., 2019; Krogh-Jespersen, Kaldy, Valadez, Carter, & Woodward, 2018]. In both cases, the presence of kinematic information allowed the ASD group to make successful predictions, but its removal led to impaired prediction, suggesting that kinematic or temporal extrapolation may not be specifically impaired in ASD (although it is worth noting that these experiments were not designed to push this predictive skill to its limit).

Finally, Deschrijver, Bardi, Wiersema, and Brass [2016] found that the ASD group's response time was more deeply modulated by violations of their expectation of object permanence, although they found a trend in the other direction when these expectations were acquired through taking the perspective of another agent. This provides further evidence that predictions based on continuing motion or stationarity are not impaired and may in fact be stronger in ASD than in NT individuals.

Social expectations. Four studies looked at predictions based primarily on social expectations, which are presumably learned over a lifetime of social experiences but may also have an innate component [Wynn, 2008]. All reported some type of impairment in making such predictions: Chambon et al. [2017] and von der Lühe et al. [2016] found that social cues induced weaker perceptual biases in action identification in the ASD group; Palumbo, Burnett, and Jellema [2015] found that biases in identification of emotions in changing faces were driven by pattern extrapolation rather than emotion-related priors in the ASD group; and Barzy, Black, Williams, and Ferguson [2019] reported differences in eye movement (though not in EEG activity), indicating differences in predictions based on voice/topic associations.

Learned statistics. *Associations learned during a single exposure.* Two studies considered predictions based on a single exposure to an antecedent-consequence pairing. Northrup, Libertus, and Iverson [2017] found that a single exposure to a noisy rattle was not sufficient to induce rattle-shaking in infants at high-risk for developing ASD when they were given a second identical rattle, indicating that at this age a single exposure may not be sufficient to create a predictive model. Krogh-Jespersen et al. [2018] found that the spontaneous gaze of young children with ASD predicted that actors would repeat the movement they made in the first exposure, but not that they would aim for the same goal. This result indicates an ability to make predictions based on a single exposure, but a tendency to make these predictions based on kinematic factors rather than implied mental states of the actor. Importantly, since these predictions were based on a single exposure, the difference between the ASD and NT

groups reflects a difference in bias in learning causal relationships rather than a difficulty in extracting statistical information.

Accumulated statistics without clear antecedent. Seven studies investigated predictions based on statistical regularities learned over the course of the experiment. In these studies, there was no clear antecedent for prediction: participants learned what stimulus qualities to expect for each trial based on the distribution of the previous stimuli. The neural or behavioral responses of NT participants were generally optimized for the arrival of a stimulus that was similar to the mean or mode of the previous ones.

A common experimental design that can be understood as investigating prediction is the "oddball" paradigm, in which a repeated stimulus is occasionally replaced by a deviant stimulus and the brain response to this predictive violation is studied using EEG. In its ordinary form, this paradigm is mainly suited to investigating predictions (and the effect of violations of those predictions) based on the expectation of repetition. However, by varying deviant probability, this paradigm can be refocused on the influence of environmental statistics on those predictions. Goris et al. [2018] found that in the ASD group, MMN responses to deviants were less modulated by the frequency of those deviants. Westerfield, Zinni, Vo, and Townsend [2015] found that the later P3 response was similarly less sensitive to global probability when the deviant was task-irrelevant, but that it was *more* sensitive to global probability when it was the target of a behavioral response task.

Three other studies looking at predictions based on accumulated statistics were perceptual studies. Two showed a reduced influence of accumulated stimulus statistics on perceptual report [Karaminis et al., 2016; Lieder et al., 2019]. Lieder et al. [2019] further broke down this influence into the influence of more recent and less recent trials and found that the ASD group's reduced use of trial statistics could be better explained by a decreased influence of recent trials. The third [Chambon et al., 2017] showed ordinary learning and biasing application of stimulus statistics on perceptual report; this was contrasted against the lack of social bias in the ASD group in a social version of the task.

Finally, two studies used eye tracking to study action anticipation based on statistical regularities in the behavior of a goal-oriented actor (i.e., repeated action toward the same goal) [Ganglmayer, Schuwerk, Sodian, & Paulus, 2020; Schuwerk, Sodian, & Paulus, 2016]. Both showed a reduced use of these repetitions to make predictive eye movements in the ASD group. These results are consistent with the hypothesis that goal information was not immediately taken to be salient or task-relevant by the ASD group, slowing learning of a predictive model.

Statistical associations learned over multiple exposures. Seven studies investigated predictions based primarily on associations between a variable antecedent or context and a consequence learned over the course of multiple exposures. Of these, one reported unambiguously behaviorally impaired prediction: ASD children's perceptual inferences indicated that they did not learn to predict an action from a statistically associated context in which that action occurred [Amoruso et al., 2019]. Lawson, Mathys, and Rees [2017] found no overall impairment in learning tone/image associations, but learning differences were observed in differential modulation of learning rate in conditions of rapidly changing associations ("volatility"). In a similar experiment with contrasting results, Manning, Kilner, Neil, Karaminis, and Pellicano [2017] found no intergroup differences in the process of learning and relearning color/reward associations. Two studies found that groups of ASD children were superior to NT children in learning associations between time interval durations and a binary outcome, as indicated by RT [Kunchulia, Tatishvili, Lomidze, Parkosadze, & Thomaschke, 2017; Kunchulia, Tatishvili, Parkosadze, Lomidze, & Thomaschke, 2020]. This evidence indicates that there is no general behavioral impairment in antecedent/consequence association learning in ASD, but there may be learning differences that depend on the value and salience of the association. Note that in Amoruso et al. [2019], the antecedent context was not necessary for performing the task during learning, and in Lawson et al. [2017] the antecedents were separated from the consequents in time and sometimes had no predictive value; whereas in the three other studies, antecedents consistently had useful predictive value and abutted the consequence in time. We elaborate on this idea further in the Discussion.

The two remaining studies in this category exposed participants repeatedly to simple associations and found notable differences in fMRI correlates of prediction errors based on such associations after they were learned. Utzerath et al. [2018] found that the ASD group showed a trend toward a less suppressed visual cortex response on seeing expected task-irrelevant pairings, indicating reduced use of task-irrelevant predictive information. By contrast, Mosner et al. [2019] found increased activity in various frontostriatal regions in the ASD group in response to task-relevant reward prediction errors, indicating increased use of task-relevant predictive information. This increase of activity associated with salient and relevant stimuli, but decrease in activation associated with task-irrelevant pairings, is consistent with a reduction in predictions based on associations that are task-irrelevant during learning.

Statistical associations learned over lifetime. Three studies considered predictions based on associations built through a lifetime of experience. Grisoni et al. [2019]

considered neural activity indicative of predictions based on the association between sounds and related words. The ASD group showed reduced modulation of neural activity by prediction, but only when the sounds pairings were action-related and not for natural sounds. Braukmann et al. [2018] studied predictions based on the association between objects and body parts (e.g., phone and ear) learned by infants based on their short experiential corpus. They found no differences in predictive gaze. van Laarhoven, Stekelenburg, Eussen, and Vroomen [2020] considered neural activity indicative of predictions based on the association between the sight and sound of a handclap. The ASD group showed *increased* modulation of neural activity by prediction.

Though these results are mixed, none of these studies found evidence of general impairment in predictions based on life-long experience, suggesting that a lifetime of learning of (salient) associations may be an intact basis for forming predictions in ASD.

Explicitly instructed rules. Four studies gave participants explicit information about probabilistic or deterministic stimulus contingencies (e.g., "X will follow Y 80% of the time") upon which to base their predictions. These studies found differences between the groups in fMRI activity related to social (but not nonsocial) reward prediction error [Kinard et al., 2020], fewer spontaneous anticipatory saccades toward the predictable target [Greene et al., 2019], a range of differences in predictability-related RT and EEG differences [Thillay et al., 2016], and differences in accuracy and fMRI correlates of predictions made from the perspective of another [Balsters et al., 2017].

Though this is a small number of studies with diverse methodologies, it is enough to strongly suggest that predictive differences in autism cannot be attributed entirely to difficulty in identifying antecedent-consequence associations: even when participants are explicitly informed of these associations, neural differences, differences in spontaneous anticipatory behavior, and differences in prediction through perspective-taking are evident.

Familiarity with complex stimulus. Two studies investigated prediction based on exposure to a complex stimulus, specifically a naturalistic image. One showed the participants degraded (i.e., "Mooney") images before and after intact versions of these images and measured participants' ability to use the intact exposure to help them recognize the degraded version [Van de Cruys, Vanmarcke, Van de Put, & Wagemans, 2018]. This study found no evidence of impairment in ASD. The other used a similar approach, but measured participant eye movement [Król & Król, 2019]. The ASD group showed less optimization and repetition of eye scan path after exposure to the intact image. The difference between these results indicates that ASD participants are able to use single perceptual experiences to influence and optimize their subsequent perceptual processing, but that the influence of these experiences on their spontaneous eye movement patterns is qualitatively different, possibly reflecting less predictive use of eye movement to optimize perceptual information uptake.

Self-initiation. van Laarhoven, Stekelenburg, Eussen, and Vroomen [2019] looked at predictions based on self-action, as measured by evoked EEG response attenuation. The evoked responses of the ASD group were less attenuated by self-generation than those of the NT group.

Results Organized by Evidence of Prediction

The 47 included articles were also divided into fourteen broad categories of predictive processes according to the evidence of prediction provided in each study.

Perceptual priors. Seven articles had participants provide responses reflecting their perceptual experiences and observed prediction by quantifying its effect as a "prior" that biased responses toward expected outcomes. In three studies, participants made inferences based on perceptual cues, and these inferences were potentially biased by prior expectations [Amoruso et al., 2019; Chambon et al., 2017; von der Lühe et al., 2016]. In the remaining studies, participants made difficult perceptual judgements that were potentially biased by prior expectations [Karaminis et al., 2016; Lieder et al., 2019; Palumbo et al., 2015; Van de Cruys, Vanmarcke, Van de Put, & Wagemans, 2018].

All but one (Van de Cruys et al., 2018) of these studies reported that responses from the ASD group were less influenced by priors than responses from the control group [Amoruso et al., 2019; Chambon et al., 2017; Karaminis et al., 2016; Lieder et al., 2019; Palumbo et al., 2015; von der Lühe et al., 2016]. The bases for these priors included social factors, associations learned over trials, and statistics accumulated over trials.

Van de Cruys et al. (2018) found no significant difference in perceptual bias between the ASD and NT group; single exposures to images helped recognition of degraded versions of these images in the ASD group as well as the NT group. This study differed from the others in that the priors were learned through a single exposure; thus, this dissenting result could suggest that the way in which the prediction or prior is learned is an important factor in whether it will be affected by ASD, and that predictions that need to be learned over multiple exposures may be more impacted.

Perceptual adaptation. Four studies [Karaminis et al., 2020; Turi et al., 2015, 2016; van Boxtel

et al., 2016] looked at the biasing influence of recent stimulus history on perception. In these studies, recent stimuli influence participants' perceptual reports by leading them to believe that the next stimulus was more different from its predecessor than it actually was. One way to interpret this effect is to say that the participant predicts that the next stimulus will be similar to the previous one, and uses that prediction to alter the neural representation of the next stimulus to improve their perceptual discrimination performance.

Three of these studies showed reduced perceptual adaptation in the ASD group [Turi et al., 2015, 2016; van Boxtel et al., 2016]. The fourth found no evidence of reduced adaptation of judgements of biological motion [Karaminis et al., 2020], in apparent contradiction to the results of van Boxtel et al. [2016]. The difference between these results may have been attributable to attentional differences, which, as shown by post hoc analysis of eye tracking data, had a significant impact on perceptual adaptation.

Response time. Seven studies had participants make simple perceptual judgments or rapidly perform other tasks in response to stimuli. The RTs indicated their readiness for each specific stimulus and/or response, which indicated the extent to which they had predicted that stimulus and/or response. [Barzy et al., 2019; Deschrijver et al., 2016; Kunchulia et al., 2017, 2020; Lawson et al., 2017; Tam et al., 2017; Thillay et al., 2016].

The results of these studies were mixed and nuanced. Lawson et al. [2017] and Thillay et al. [2016] both found less modulation of reaction time by stimulus predictability in ASD. However, in Lawson et al. [2017] this result became more ambiguous when the results were corrected using computational models to account for differences in learning the predictive associations. And in Thillay et al. [2016], this difference was due not to slowed RTs in sequence-based predictable contexts but instead to faster RTs in unpredictable contexts, indicating strong motor preparation that was less sensitive to target probability (and, interestingly, lacking in motor-preparatory desynchronization of the EEG mu rhythm in motor areas). Tam et al. [2017] found a lack of improvement in reaction time over repetitions of a working memory task, which they argued represented a failure to learn to expect the distracting stimuli in the task.

Of the remaining four studies, one showed no significant difference in modulation of RT by predictability [Barzy et al., 2019]. The remaining three found *increased* modulation of RT by predictability in the ASD group. Kunchulia et al. [2017] and [2020] found a greater decrease in RT when the outcome was predictable due to learned interval duration cues. Deschrijver et al. [2016] found a greater increase in RT when the outcome was unpredictable due to violation of an expectation of object permanence, though RT trended weakly in the other direction when these expectations were based on observations from the perspective of an actor in the scene.

Together, these results indicate no systematic impairment of predictive modulation of response time. Rather, results appear to vary with the type of associations on which the predictions are based. However, the findings of Thillay et al. [2016] showing differences in RT and its EEG correlates even when predictions are based on explicit instructions indicate that the story of prediction-modulated RT in ASD is multifaceted and requires additional study.

Reward-oriented decision-making. Manning et al. [2017] had participants indicate their predictions by choosing which of two options they thought would maximize their expected reward. The authors found no significant differences in performance or strategy between the ASD and NT children on their task, indicating no apparent alterations in expressing their predictions through this type of decision-making.

Gaze. Ten studies had participants indicate their predictions through oculo-motor engagement. Generally, the participants had explicit or implicit goals in the task, and their oculo-motor behavior indicated the extent and accuracy of their use of predictions to achieve that goal.

Eight of these studies measured anticipatory eye movement as an indicator of prediction [Barzy et al., 2019; Braukmann et al., 2018; Ego et al., 2016; Ganglmayer et al., 2020; Greene et al., 2019; Krogh-Jespersen et al., 2018; Schuwerk et al., 2016; Vivanti et al., 2018]. While two studies found no differences in predictive eye movements (i.e., equal amounts of predictive saccades based on common action/body part associations [Braukmann et al., 2018] and comparable predictive tracking of a moving object through occlusion [Ego et al., 2016]), the remaining six found fewer predictive eye movements in ASD based on learned [Schuwerk et al., 2016] or instructed [Greene et al., 2019] statistical tendencies. Anticipatory saccades in ASD participants were either executed with a greater delay after the predictive cue [Barzy et al., 2019] or not similarly informed by the inferred goals of agents in the scene [Ganglmayer et al., 2020; Krogh-Jespersen et al., 2018]. Differences in the number of anticipatory saccades may have been attributable to the lack of a visible target in the two experiments reporting fewer saccades [Goldberg et al., 2002].

Studies looking at predictive eye movements based on kinematic information found them unaffected [Ego et al., 2016; Krogh-Jespersen et al., 2018]. Vivanti et al. [2018] used eye fixation as an indicator of the habituation of attention by repetition of visual stimuli, which could arguably be interpreted as a measure of how quickly observers began to predict the repetition. They

found that the ASD group continued to fixate on a repetitive visual stream for longer than the NT group. Król and Król [2019] found that the ASD group did not use repeated exposure to degraded images to optimize their scan path for recognition to the same degree as the NT group.

These results indicate that individuals with ASD exhibit some atypicalities in predictive eye movements. Such movements may be reduced when they are directed toward locations without visible targets. Additionally, atypicalities are modulated by the bases for prediction; statistical and learned goal-based predictions are particularly effective in eliciting differences while basic kinematic predictions are less so.

Overall, evidence of compromised predictive eye movements could be interpreted as a systematic difference in making or in learning to make eye movements informed by the predictability of stimulus. Results from Ego et al. [2016] support the latter interpretation: when participants were instructed in what anticipatory eye movements to make, no difference was found between groups.

Movement. Two studies measured spontaneous anticipatory motor behavior in infants at risk of ASD as an indicator of their predictions [Landa et al., 2016; Northrup, 2017] and found these behaviors under expressed. Landa et al. [2016] found that infants at high risk for ASD less frequently reached to intercept a rolling ball. Northrup et al. [2017] found that infants at high risk for ASD started shaking rattles later, possibly indicating reduced anticipation of feedback.

Two further studies had participants engage in deliberately coordinated real or virtual movement and used associated parameters (e.g., timing, position, speed) as measures of prediction [Fulceri et al., 2018; Tewolde et al., 2018]. Tewolde et al. [2018] found that the ASD group was similarly able to anticipate an arrival time based on temporal and kinematic extrapolation, whereas Fulceri et al. [2018] found that the ASD group had greater difficulty coordinating a joint action when given only kinematic cues from which to extrapolate. This discrepancy might have been attributable to the first but not the second task having a clear behavioral goal that the participants could focus on, and suggests that individuals with ASD may be quite capable of coordinating instructed anticipatory movements but may not use anticipation to optimize their spontaneous movements.

Pupillometry. Lawson et al. [2017] looked at pupil diameter upon observation of the consequence as evidence of whether it was predicted. This is a physiological response that presumably provides an evolutionary advantage for the organism by optimizing its vision for expected vs. unexpected situations. They found reduced modulation of pupil dilation by predictability in the ASD group; however, they found increased modulation of pupil dilation by a model-based measure of surprise that took into account differences in learning of predictive associations. This result highlights the value (and difficulty) of separating out the effect of differences in learning underlying predictive relationships from low-level differences in prediction-modulated responses.

Electrical activity in anticipation of a predictable consequence. Two studies analyzed EEG activity prior to the onset of predictable stimuli as a measure of expectation. Individuals with ASD were observed to show similar ramping activity (contingent negative variation, or CNV) preceding predictable targets as well as non-targets, while TD controls showed a greater CNV amplitude preceding predictable targets only [Thillay et al., 2016]. This suggests that individuals with ASD are not efficiently using prior information to make predictions about upcoming events. Similarly, Grisoni et al. [2019] found a reduction of predictive potential (PP) in individuals with ASD, compared to controls, when they were presented with pairs of predictable words and sounds that related to actions.

Electrical activity in response to a predictable consequence. As it relates to predictive impairments, EEG has primarily been used to study the electrical activity associated with a response to predictable stimuli. A repeated, predictable stimulus typically causes an attenuation of early peaks (i.e., P1, N1, P2) within the EEG waveform, which may be evidence of low-level predictions of incoming stimuli. This habituation may be related to the conservation of neural resources important for optimizing one's response to unexpected input. It is hypothesized that a reduction of this habituation may reflect atypical predictive processing, and may partially explain the sensory hypersensitivities associated with ASD [Sinha et al., 2014]. Further, the unexpected omission of a stimulus results in an early omission evoked response (omission N1; oN1).

Ruiz-Martínez et al. [2020] and van Laarhoven et al. [2019] used attenuation of the early component of EEG stimulus-evoked activity as a measure of prediction. Both showed reduced modulation by prediction of the early ERP response to auditory stimuli in the ASD group compared to the control group. A third study examining the effect of auditory omission during an auditory visual action (i.e., hand clap) showed an increased oN1 in the ASD group compared to the NT group, suggesting a larger prediction error in the ASD group compared to the NT group [van Laarhoven et al., 2020].

Later components of the EEG waveform (N2, P3, N4), and second order EEG measures such as mismatch negativity (MMN), are sensitive to the perceived likelihood of an upcoming stimulus. The amplitudes of N2, P3, N4 are inversely correlated with predictability in neurotypical

subjects, and thus can be interpreted as measures of novelty detection or of prediction. Similarly, the amplitude of the derived MMN, which is calculated by subtracting an averaged waveform in response to standard stimuli from an averaged waveform in response to a deviant stimulus, is also a measure of novelty or change detection. As such, experimenters can modulate the probability of a stimulus and use the amplitude and latency of these components as a measure of prediction. Eight studies analyzed one or more later or derived ERP components as a measure of prediction. All but one study found differences in these prediction-based signals in the ASD group [Gonzalez-Gadea et al., 2015; Goris et al., 2018; Grisoni et al., 2019; Ruiz-Martínez et al., 2020]. One study demonstrated that changes to the P3 amplitude as a function of predictability were attention dependent in ASD [Westerfield et al., 2015]. The one study that did not find group differences in these prediction-based signals examined N4 amplitude as a function of linguistic predictability based on speaker identity [Barzy et al., 2019].

Overall, the majority of EEG-based evidence generally indicates a predictive difference in ASD, though because there is not a clean correspondence between these EEG signals and behavior, it is difficult to draw conclusions about functional behavioral differences in prediction from these data.

BOLD signal in response to a predictable consequence. Four studies analyzed fMRI blood-oxygen-level-dependent (BOLD) signals in response to a repeated stimulus. As described above, neural habituation, also referred to as adaptation or repetition suppression, to a repeated stimulus over time can serve as a measure of low-level prediction. Three of the fMRI studies revealed a reduction of habituation in the ASD participants compared to controls [Ewbank et al., 2017; Millin et al., 2018; Tam et al., 2017]. These results were nuanced: individuals with ASD demonstrated impaired habituation to auditory but not visual stimuli [Millin et al., 2018], and to faces (Ewbank et al., 2017; Tam et al., 2017), but not shapes (Ewbank et al., 2017). One study demonstrated intact repetition suppression to shapes in the visual cortex, but a difference in response to expected compared to unexpected faces in V1 [Utzerath et al., 2018].

Three studies analyzed BOLD signals related to reward prediction errors (RPE). One study used non-social learned outcome expectancies [Mosner et al., 2019], and two used social conditions [Balsters et al., 2017; Kinard et al., 2020]. The social conditions within these studies were quite different; they included stimuli that were social (smiling face) and non-social (preferred objects) [Kinard et al., 2020] and a theory of mind task with first and third person perspective taking [Balsters et al., 2017]. All three studies showed differences in the frontostriatal prediction error signal between the ASD group and the NT group. One study found the neural differences during social, but not non-social, reward conditions for adolescents with ASD [Kinard et al., 2020].

Reported prediction or surprise. Balsters et al. [2017] had participants indicate their predictions by reporting whether they perceived a reward outcome to be expected or unexpected. They found that the ASD group more frequently reported the likelier event as unexpected (and vice versa), especially when describing the expectedness of rewards given to others and not themselves. A second study had participants indicate their predictions by explicitly reporting what they thought would happen, specifically which of two cars they thought would arrive first [Sheppard et al., 2016]. They found that the ASD group gave less accurate predictions, but only when there were limited self-motion cues. Both studies found predictive differences between the ASD and NT groups.

Discussion

Since the proposal of the PIA hypothesis in 2014, there has been considerable theoretical discussion and empirical research investigating the veracity of this theory. In this review, we have attempted to distill the empirical study of prediction into two main observable components of prediction: the basis for prediction and the evidence of prediction. These two organizational schemes allow us to look separately for possible differences in the predictive processes used by individuals with ASD and for possible differences in their ability and tendency to form the predictive associations that underlie the predictive processes.

The organizational scheme (i.e., bases for prediction and evidence of prediction) that we have used for this review fits within the structure provided by two prominent theoretical papers on prediction impairments in ASD. It is helpful to orient our literature review, and the multiple hypotheses motivating it, along an axis from *predictive process impairment* to *predictive learning impairment* (Fig. 3). Although the two perspectives are richly linked, especially within the context of the predictive processing formulation of cognition, they offer landmarks by which to navigate a complex dialogue.

Predictive Learning Impairment

A proposal of *predictive learning impairment* was put forward by Sinha et al. [2014], who argued that ASD was characterized by inaccuracies in learning the probabilities $P(B|A, \Delta t)$ that govern transitions in external events over time, probabilities necessary to make accurate predictions with an appropriate degree of certainty. This perspective builds on a literature exploring differences in implicit

and statistical learning in ASD. These differences are difficult to pin down: various studies of implicit and statistical learning have shown no behavioral differences in ASD [Foti, De Crescenzo, Vivanti, Menghini, & Vicari, 2015; Mayo & Eigsti, 2012; Obeid, Brooks, Powers, Gillespie-Lynch, & Lum, 2016], but EEG and brain imaging evidence suggests that participants with ASD may be using a different and perhaps more intentional learning strategy to succeed in these tasks [Scott-Van Zeeland et al., 2010; Zwart, Vissers, van der Meij, Kessels, & Maes, 2017].

In this literature review, we found mixed evidence of differences in learning predictive associations between antecedent and consequence, which we will attempt to assemble into a coherent picture of the conditions under which predictive learning is impaired in ASD.

Limited scope of contingency learning. In the category of predictions based on repeated exposure to antecedent-consequence pairings, successfully learned predictive pairings tended to be those with antecedents that had consistent, task-relevant predictive value during learning and temporally abutted or overlapped with the consequence. Pairings that did not fit this description showed evidence of impaired learning. Although these two factors—inconsistent predictive value and temporal separation -may represent two different difficulties in learning predictive associations, we cautiously group them together as situations in which the antecedent's predictive connection to the consequence may be less detectable or less apparently relevant (i.e., have less *predictive salience*).

An impairment in detecting and learning predictive associations may be a manifestation of *stimulus overselectivity*, as originally identified and then reviewed by Lovaas, Koegel, and Schreibman [1979] and more recently reviewed by Ploog [2010]. This refers to autistic individuals' tendency to learn to associate only one perceptual aspect of a complex predictive cue with its predicted affordance of a reward. Which aspect is prioritized may vary from individual to individual: in early studies, some participants learned only the visual component of a multimodal cue while others learned only the auditory component. Thus, it may be useful to focus on the narrower bandwidth for learning in ASD than on any specific stimulus features that may or may not be used for learning. A narrowed scope for learning antecedents may have the effect of limiting learning to only those associations where the antecedent is expected a priori to have potential value in predicting the consequence.

An impairment in detecting subtler contingencies could plausibly underlie the consistent predictive impairments in studies in which predictions depended on social priors. In Northrup et al. [2017], the author describes a theory of social development centered around the detection and exploitation of *contingencies* (what we are calling *predictive associations*). In this model, social contingencies are learned through a cycle of learned patterns of attention and positive experiences with social prediction. This is the point at which underlying differences in contingency detection in infants at high risk for ASD are magnified by a cascade of learning effects that cause their social development to diverge from typical development.

Several other studies with predictions based on goal-learning can be explained in terms of difficulties with contingency learning and resulting irregular development of attunement to social contingencies. In both Krogh-Jespersen et al. [2018] and Ganglmayer et al. [2020], effective prediction of an agent's action depended on observing and learning its goals based on its preceding actions. An arrested cycle of social contingency learning could lead to a brain that is less attuned to agent goal as a highly salient aspect of an observed action. Coupled with a learning process that is largely restricted to the most salient aspects of a stimulus, this could explain the slowed or absent learning of goal-based predictions in these studies.

The hypothesis of slowed association learning due to differences in antecedent salience is buttressed by conditioning results not included in our systematic review. In a rat model of ASD, Kosaki and Watanabe [2016] showed that Pavlovian conditioning was ineffective at connecting audiovisual antecedent stimuli with rewarding or aversive consequences. These rats could, however, learn predictive relationships when the antecedent overlapped the consequent in time and was either immersive (one room vs. another) or in a modality already connected with the consequence (taste antecedent, gastric distress consequence). The authors describe this result as indicating "dysfunctional contingency-based, but not contiguity-based, learning," where contingency-based learning requires selectively identifying a predictive antecedent among all the sensory stimuli present before the consequence. Similarly, in a study of fear conditioning in humans, the ASD group was less receptive to conditioning [Powell, Travers, Klinger, & Klinger, 2016]. In this case, ASD participants with a greater explicit awareness of the conditioning contingency showed more conditioning, consistent with the hypothesis of a narrower bandwidth for contingency learning limited to more salient aspects of the antecedent.

Such a predictive learning difference could be understood in the context of learning the successor representation (SR), a construct created to explain phenomena in reinforcement learning but which has broad implications [Dayan, 1993; Gershman, 2018]. The SR is a learned estimate of the tendency of any state to eventually lead to any other state based on a set of sensory features of the states and discounted by transition time. According to Gardner, Schoenbaum, and Gershman [2018], this representation is likely stored in the hippocampus and learned by *dopaminergic signals of prediction error*. One dopamine channel is necessary for each sensory feature. Learning

differences in ASD could be due to differences in dopaminergic signaling or differences upstream at the level of attention, as discussed below.

Predictive Process Impairment

As shown on the process impairment end of Figure 3, the main *predictive process impairment* hypothesized as a unifying feature of ASD is the *hypo-prior* hypothesis introduced by Pellicano and Burr [2012]. These authors brought together a rich history of research on perceptual differences in ASD [Mitchell & Ropar, 2004] and a growing literature describing perception as a process of Bayesian inference in which data are assimilated in the context of prior expectations. Pellicano and Burr [2012] hypothesized that the brain of an individual with ASD draws less heavily on priors in its perceptual inferences, creating a perceptual experience more loyal to and more bound by current sensoria. This account located the key predictive difference in ASD as a difference in the perceptual manifestations of prediction (though it left open the possibility that the learning of priors was impaired rather than their perceptual influence). In response, Brock [2012] pointed out that the important variable in Bayesian perception is the weight of the prior *relative to the weight of sensory data*, and thus that the characteristic differences in autistic perception could be equally well described by positing an overweighting of sensation.

Below, we consider three areas of our literature review in which we find accumulations of evidence suggesting process impairments.

Prediction as perceptual prior. The seven studies since 2014 showing reduced influence of perceptual priors (reviewed in Section 3.2.1) provide weighty evidence of an impairment along the lines proposed by Pellicano and Burr [2012]. However, a second look through this evidence suggests that the difference in ASD might be more closely related to learning what to predict rather than using that prediction as a prior to bias perception. In these studies, the priors were all either "empirical priors" (learned over sequential exposures during the experiment) or "systemic priors" specific to the social domain. Weak priors in the latter group could plausibly be attributed to an impairment in lifelong social learning. Further, one of these studies showed that weak social priors could be compensated for by empirical learning over trials, indicating that priors could be effectively employed by the ASD group once successfully internalized [Chambon et al., 2017]. Finally, the one dissenting study showed that single-exposure perceptual learning is equally helpful for distorted image recognition among ASD participants (Van de Cruys et al., 2018). These results, combined with results not reviewed here indicating intactness of static priors (Van de Cruys et al., 2018) (perceptual biases not dependent on an antecedent context but applicable throughout life), indicate that ASD is not characterized by an impairment in employing predictions as priors to bias perception, but instead by differences in the process of learning the predictive models that form the basis for these instances of biased perception, as discussed above. Further, Van de Cruys et al. (2018) is one of the few studies in this set where it is clear throughout the experiment that perceptual history will be relevant for successful performance of the task at hand; thus, this result could be taken to support the role of task-relevance in learning predictions.

Reduced spontaneous motor prediction as evidence of process impairment. The literature reviewed here showed consistent differences between ASD and NT groups in spontaneous anticipatory gaze. These differences spanned multiple bases of prediction, including predictions based on explicit instructions, and are therefore not easily attributed to difficulties learning what to predict. These results may be related to finding from the two studies looking at spontaneous anticipatory arm movement in infants who were at high risk for ASD, which both reported reduced predictive movement in the ASD group, and to the reduced EEG signatures of predictive motor preparation shown in Thillay et al. [2016]. However, deliberate predictive movement seems to be unaffected by ASD: in several studies, participants showed intact predictive eye movements and motor engagement when they were explicitly instructed to move predictively. Taken together, this literature suggests a general underutilization of spontaneous predictive motor processes, a conclusion in line with earlier results suggesting reduced predictive feeding behavior in children and infants with ASD [Brisson, Warreyn, Serres, Foussier, & Adrien-Louis, 2012; Cattaneo et al., 2007].

However, such an impairment could be related to learning differences of the predictive behavior, rather than in learning the underlying predictive associations. According to the theory of active inference for motor control (a part of the predictive processing framework), motor plans are first formulated as predictions of proprioceptive stimuli and then manifest as movement by the lowest levels of the motor system. Thus, learning to move in anticipation may require learning to expect yourself to move in a certain way in a certain situation, and the rate and breadth of learning may be limited in the same way as we have described above for the prediction of external stimuli.

Prediction based on stimulus constancy. Overall, the results of our review suggest reduced habituation of neural activity and behavior (gaze and response time) over stimulus repetitions in the ASD group compared to the NT controls. Habituation studies demonstrate impaired auditory

habituation in the ASD group. Interestingly, in the visual domain, habitation to faces appears to be uniquely impaired. This face-specific habituation result has been replicated by other researchers [Kleinhans, Richards, Greenson, Dawson, & Aylward, 2016; Webb et al., 2010]; however, only one of these studies met the inclusion criteria for this review. The results further show reduced perceptual adaptation, which can be interpreted as reduced use of the expectation of stimulus repetition to improve stimulus discrimination. However, one study indicated that this effect may be attention-dependent [Karaminis et al., 2020], as we discuss below.

The prediction that a stimulus will be similar to the preceding stimulus is not static, in that a second stimulus is only predicted once there has been a first one, but predictions of this form are applicable in all settings and in all modalities. The tendency for stimuli to repeat is a global predictive constraint, and is therefore likely to be implemented through bottom-up mechanisms at earlier stages of sensory processing along multiple separate sensory streams [Teufel & Fletcher, 2020]. It is also possible that the processing of repeating stimuli occurs through separate top-down and bottom-up mechanisms along the same sensory stream [Ewbank, Henson, Rowe, Stoyanova, & Calder, 2013]. Intact mismatch-negativity responses to vowel sounds in pre-term infants [Cheour et al., 1998] suggest that at least some sensory streams have acquired such an implicit expectation of repetition at birth.

The very early emergence of the expectation of repetition, and the low level of the sensory stream at which the learned component of these expectations likely takes place, makes it difficult to conceptually distinguish between a possible failure to learn the expectation of repetition and the failure to use it to initiate a predictive process. Neural circuit differences causing, for example, reduced habituation to repetition could be interpreted as the mechanistic underpinning of a failure to use the repeating pattern to initiate a predictive process that can attenuate the sensory signal, or as a failure to learn the repeating pattern with the neural circuits necessary to use it in a predictive manner.

Precision Modulation Impairment

The past 20 years have seen a compelling attempt to unify many of the manifestations of prediction (as well as the rest of cognition) under the banner of *predictive coding* or *predictive processing* [Friston, 2005, 2018; Rao & Ballard, 1999]. According to the predictive coding framework, the brain's activity and connectivity functions as a hierarchical model of processes and hidden causes in the world that give rise to sensory data. A *prediction* is the sensory input expected by the brain given its estimates of hidden states. These predictions may be more or less *precise* (i.e., certain). All change within the brain plays out according to a mandate to minimize the sum of the *errors* by which the predictions deviate from real sensory data, with more precise predictions being given higher weight than less precise predictions.

Importantly, this framework describes both changes in ongoing electrical activity and changes in connectivity between neurons. At short time scales observable through EEG and fMRI, brain activity representing hidden states in the world is adjusted in response to every observation to minimize the error in predictions that arise from it, and predictive processes in neural activity and behavior play out in order to minimize future prediction errors. At longer time scales, error-minimizing adjustment in connection strength constitutes our learning about the rules governing our sensations. Note that although both of these impacts of prediction follow the same mandate of minimizing precision-weighted prediction error, they may be implemented through different neurophysiological mechanisms, and therefore may be differently affected by neurological variability and disorder.

After the influential Pellicano and Burr [2012] paper on impaired priors in ASD, several groups of researchers quickly responded to recast the discussion into the "predictive coding" framework [Friston, Lawson, & Frith, 2013; van Boxtel & Lu, 2013; Van de Cruys, de Wit, Evers, Boets, & Wagemans, 2013]. Soon afterward, two of these groups fleshed out predictive impairment hypotheses centered around the handling of predictive and sensory precision. Lawson et al. [2014] offered an account of autistic symptoms largely aligned with the notion of process impairment, hypothesizing a "failure to attenuate sensory precision and contextualize sensory information in an optimal fashion," while acknowledging that "statistically 'optimal' perception and learning go hand in hand" (Fig. 3). Van de Cruys et al. [2014] provided an account closer to the learning-based proposal, claiming that "predictive coding can much more naturally explain the range of changes seen in autism by focusing on how priors develop on the basis of prediction error." In this account, the central impairment in ASD is in high, inflexible precision of prediction errors, which force the learning of new predictions even in contexts that are inherently unpredictable.

Shared by both of these accounts is the view that ASD involves not a uniform imbalance in precision but a failure to *contextually modulate precision*, resulting in either perception that is overly veridical or learning that is overfit for the context. They discuss two classes of contextual modulation of precision:

1. Modulation as an immediate result of making prediction errors (*reactive* modulation in response to unexpected uncertainty), thought to be mediated by norepinephrine [Yu & Dayan, 2005].
2. Modulation based on context clues that are associated with high or low (prior or sensory) precision through learning (*predictive* modulation in preparation for

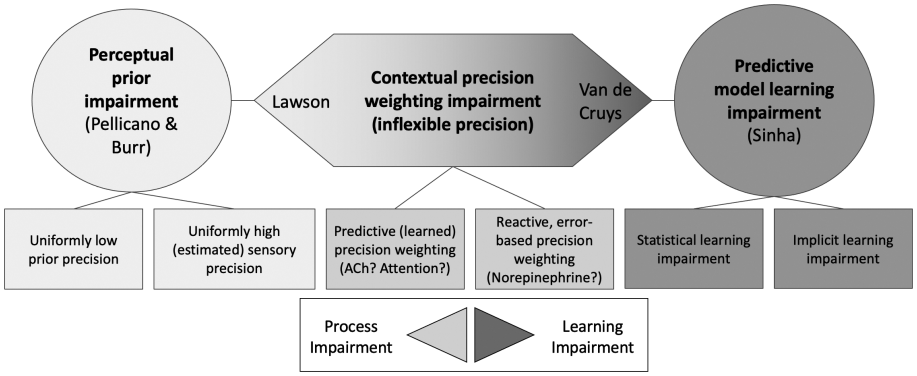

Figure 3. Visualization of the prominent theories of prediction in ASD.

expected uncertainty), thought to be mediated by acetylcholine [Yu & Dayan, 2005].

Evidence has linked autism to possible impairments in both neuromodulatory systems, providing a possible neurophysiological mechanism for predictive processing differences in ASD.

Prediction error signaling as precision modulation impairment. One way to understand the differences between ASD and NT groups in neural evidence of prediction based on stimulus constancy—specifically, differences in neural habituation—may be in terms of prediction error signaling. Since these predictions are either innate or learned extensively and early, it is probably safe to assume that the ground truth predictions are in place, eliminating the learning differences discussed above as a possible source of differences in neural activity. In the Van de Cruys precision impairment hypothesis, it is the lack of predictive attenuation of neural responses that gives rise to a cascade of higher-level effects in perceptual and learning processes that rely on these neural signals of prediction error.

In the reviewed literature, the early component of ERPs showed consistent evidence of reduced habituation to tones in the ASD group that could indicate less predictive attenuation of sensory input than in the NT group. A lack of behavioral (gaze and reaction time) attenuation could represent direct results of this signaling difference. Another prediction type likely to be innate or strongly learned in the ASD brain is the prediction of a stimulus when it is self-initiated, so the observation of relatively attenuated neural response to self-generated tones supports this hypothesis. Finally, the presence of a stronger omission response to the sight of a handclap without the sound [van Laarhoven et al., 2020] suggests that predictive processes based on strongly learned associations are present in ASD, and that signaling differences may stem from an over signaling of prediction errors rather than a failure to predict.

Reactive precision modulation as precision modulation impairment. Notwithstanding the focus of several prominent ASD hypotheses on prospective and reactive contextual modulation of predictive precision, literature directly addressing this is scarce. This is likely due in part to the difficulty of sorting out the interaction of aberrant prediction error signaling and atypical learning of the predictable ground truth. The two reviewed studies that directly investigated reactive precision modulation—the capacity to modulate error signaling and learning rate based on patterns of prediction errors—dealt with this problem by describing behavior with normative computational models and fitting parameters representing aspects of learning and of prediction-modulated response to groups or individuals [Lawson et al., 2017; Manning et al., 2017]. The studies gave conflicting results, as we have discussed previously. Their approach is promising but requires careful navigation. Suitably nuanced models have many free parameters, and interpretation of these parameters is only as strong as the assumptions underlying the model.

Evidence of precision impairment in ASD: The effect of attention. The critical role of attention in prediction lurks just below the surface in the reviewed literature, emerging repeatedly as a possible explanation for unexpected effects. However, if the key predictive impairment in ASD is indeed modulation of prediction precision, attention should be central to the discussion: within predictive processing, attention is understood as a mechanism for precision modulation, with top-down attention corresponding to prospective precision modulation based on inferred context [Feldman & Friston, 2010], likely moderated by acetylcholine. Van de Cruys et al. [2014] attribute attentional differences in ASD to a compensatory strategy to handle overwhelming prediction error signals, but the data suggest that the role of attention in differences in autistic prediction may deserve another look.

Two studies reviewed here looked directly at attended vs. unattended predictions, with the target of attention determined by task instructions [Gonzalez-Gadea et al., 2015; Westerfield et al., 2015]. In both, neural responses were stronger in the ASD group for attended prediction errors but weaker for unattended prediction errors, suggesting that attention plays a powerful mediating role in predictive processes or prediction-modulated responses in ASD. Karaminis et al. [2020] provided evidence suggesting that perceptual adaptation in ASD participants may be normalized by attention. These results are not consistent with a uniform difficulty contextually modulating precision; rather, they suggest that top-down attention may play an oversized role in precision modulation.

Attention may play an important role in the learning differences discussed above: the narrow bandwidth for possible predictive cues that we propose above may be limited by the scope of attention. Kosaki and Watanabe [2016] link their conditioning result to a class of associative learning theories that posit that "stimulus-selective learning is achieved (at least partly) through modifications of attention paid to each stimulus on the basis of the degree with which each stimulus reliably signals the (unconditioned stimulus)" [Le Pelley, 2004; Mackintosh, 1975]. Thus, learning differences in ASD may be due to a narrower scope of attention [Robertson, Kravitz, Freyberg, Baron-Cohen, & Baker, 2013]. Alternatively, it may be that learning that can occur without explicit attention in the neurotypical brain is restricted to attended features in the autistic brain. This is consistent with the possibility of attention playing an unusually strong role in precision modulation.

Attention may even be important to prediction in the motor system. In the context of predictive coding and its extension in the motor domain (i.e., active inference), Brown, Friston, and Bestmann [2011] propose that "motor preparation is an attentional phenomenon directed toward kinesthetic signals." Thus, the complex picture of prediction differences in ASD as measured by response time may be considered through the lens of attentional dynamics.

The notion that prospective modulation of precision—adjusting error signaling and learning rates based on learned context cues—may function differently in ASD adds yet another level of complexity to the study of precision dynamics: learning the cues necessary to establish the appropriate "ground truth" level of precision. It is no wonder that a thorough study of attentional modulation of precision in ASD has not yet been undertaken. However, Mirza, Adams, Friston, and Parr [2019] have proposed a computational model of attention that attempts to account for aspects of the ASD phenotype and that might serve as a guide to further investigation of attention in ASD.

Impaired learning due to a narrowed range of salience or attentional scope is reminiscent of a range of other unifying ideas about ASD. It may be linked to the narrowed scope of spatial attention in ASD [Robertson et al., 2013], which itself may underlie the observation of weaker neural suppression [Schallmo et al., 2020]. It also resonates with the theory of more local, less global processing in ASD [Happé & Frith, 2006]. And in the predictive processing framework, attention is treated as setting a higher precision for predictions within the attended stream, and salience as a measure of precision; thus, as is pointed out in Van de Cruys, Van der Hallen, and Wagemans [2017], differences in identifying which aspects of a sensory context are salient may be one expression of the difficulty modulating precision posited by the predictive coding theories discussed above.

Exclusively Social Impairment

The prediction impairment theories of ASD need to be considered in the context of the popular conceptualization of ASD as primarily a social impairment. Although there is ample evidence for non-social predictive differences in ASD, many of the results reviewed here focus on social predictions, and some provide evidence that aspects of predictive impairment may be limited to social prediction.

Two studies looked at predictions made from the perspective of (i.e., based on antecedents observed by) another agent. In Balsters et al. [2017], ASD participants showed reduced fMRI evidence of prediction and reduced prediction accuracy relative to the NT group when their predictions were made from the perspective of another. But in Deschrijver et al. [2016], reaction times did not show a significant difference between the two groups in predictions made from the perspective of another.

Four studies found predictive differences in ASD that were specific to social stimuli and did not appear for stimuli outside the social domain: reduced neural habituation to repeated faces but not repeated shapes [Ewbank et al., 2017], reduced EEG evidence of prediction for action sounds [Grisoni et al., 2019], reduced accuracy of perceptual inferences when they were based on social priors [von Der Lühe et al., 2016], and reduced fMRI activity related to social (but not nonsocial) reward prediction error differences [Kinard et al., 2020]. However, Greene et al. [2019] observed similar differences in predictive gaze for social and nonsocial stimuli. Given the extensive evidence presented here of differences in nonsocial prediction in ASD, perhaps the few examples of predictive differences limited to social stimuli can best be understood in terms of an additional layer of predictive difference introduced by social stimuli on top of underlying nonsocial differences that may not be expressed in every task situation. Thus, these results do not entirely

undermine the proposal that social differences in ASD may only be the most prominent expression or outgrowth of underlying prediction-related differences.

Limitations

We would like to recognize that while we were systematic in our identification of articles and use of PRISMA guidelines, we may have missed articles that study behaviors and neural activity related to prediction but do not utilize the prediction terminology. For example, there are many EEG studies that analyze MMN that were not captured by our search because they have not been described within the context of the prediction literature or were published prior to the emergence of the PIA hypothesis. We are hopeful that a more consistent use of prediction-related language, described elsewhere and considered here, will make future reviews easier and further our understanding of prediction in ASD by including an even larger number of studies.

Due to the large scope of our review, and diversity of experimental paradigms included, and data collected, we did not conduct a formal meta-analysis. This is a worthy goal for future research and will be facilitated by the progressive accumulation of empirical studies that utilize shared methodologies.

We would also like to acknowledge the authors' potential biases in conducting this review. The last author wrote one of the PIA papers [Sinha et al., 2014] and all authors of this review are currently engaging in active research related to prediction in autism. In an attempt to reduce potential bias, and increase transparency, we strictly adhered to the PRISMA guidelines in conducting this review [Moher et al., 2010].

A notable limitation of the research reviewed here is that the applicability of these differences in predictive processes to real-world functional impairments is not entirely clear. The differences observed in EEG, in particular, are not clearly connected with behavioral differences. Further, the behavioral tasks that have been used are generally circumscribed and not naturalistic. This makes it more difficult to make conclusions about the daily life and behavioral impacts of the differences observed in controlled laboratory settings.

Future Directions

While there has certainly been an abundance of research related to prediction in ASD over the past six years, we have identified some areas that warrant additional attention. The studies reviewed here generally focus on one specific element or type of prediction. Given the domain-general nature of the prediction in ASD hypotheses, research that tests multiple types of prediction (e.g., sequential pattern prediction and language prediction) within the same individuals would be useful.

Naturalistic studies are also largely missing in the extant literature. Though naturalistic experiments present their own challenges, they are particularly important for understanding the daily difficulties, or lack thereof, of the observed predictive differences. In the absence of conducting studies *in situ*, lab-based studies that target prediction related to temporal and sequential patterns would also yield more translatable results to daily experiences. A self-report survey that measures prediction skills within the context of daily living may be useful for identifying when and where people experience difficulty with predictions. This type of self-assessment could then guide naturalistic experiment design and intervention. With this motivation, our lab is currently engaged in developing such a survey instrument.

Further, while many studies use socially relevant stimuli (e.g., faces, voices), there is limited research related to the use of social priors within social contexts. These studies would likely tap into more sophisticated types of prediction and be particularly relevant to ASD given the social communication deficits included in the diagnostic criteria.

While the heterogeneity of ASD symptomatology is widely recognized, the papers reviewed here did not explore the functional relationship between prediction differences and ASD symptom severity, phenotypic profiles, or genetic profiles. Future research would benefit from exploring how individual differences in prediction skills relate to ASD profiles in order to better inform our understanding of the broad ASD phenotype and customize intervention.

Beyond characterizing differences in prediction, future research should explore the benefits of providing prediction-based interventions to individuals with different styles of prediction. The reviewed research offers no conclusion as to whether predictive difficulties in ASD can be alleviated through training. However, results showing prolonged but successful predictive learning in ASD [e.g., Ganglmayer et al., 2020], and unimpaired predictive behavior under conditions of explicit instruction [e.g., Ego et al., 2016], suggest that predictive behaviors and cognitive strategies may be promising targets for intervention. If differences in predictive learning are indeed mediated by attention, then interventions that direct attention toward predictive cues may be especially useful.

One prediction-oriented intervention, a parent-mediated intervention called Predictive Parenting, was studied for feasibility and determined to be acceptable and accessible to parents [Hallett et al., 2020]. This intervention focused on teaching parents to utilize strategies that increase a child's understanding of daily activities and to help children cope with uncertainty. Other

intervention studies have focused on enhancing prediction skills by using simulators to teach new skills, including driving Cox et al., 2020].

Conclusions

Taken as a whole, the 47 studies reviewed herewith advance our understanding of the underlying differences in prediction in ASD. In particular, both the learning of predictive associations and the neural signaling of low-level prediction errors appear to be impacted in ASD. These differences likely have downstream consequences for higher-level prediction skills and may underlie social manifestations of ASD. Continued research into the nuances of prediction differences in ASD, and the resultant functional impacts for these individuals, is warranted.

Acknowledgments

We acknowledge Annie L. Cardinaux, Sidney P. Diamond, Isaac N. Treves, and Anila D'Mello for their discussions and comments during the preparation of this manuscript. This work was supported by the Simons Center for the Social Brain at Massachusetts Institute of Technology, Cambridge, MA and by the Simons Foundation for Autism Research Initiative in New York City, NY. AMO was supported by the NIH-NIDCD T32 (5T32DC000038-29).

The authors have no conflicts of interest to disclose.

References

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

Balsters, J. H., Apps, M. A. J., Bolis, D., Lehner, R., Gallagher, L., & Wenderoth, N. (2017). Disrupted prediction errors index social deficits in autism spectrum disorder. Brain: A Journal of Neurology, 140(1), 235–246.

Barzy, M., Black, J., Williams, D., & Ferguson, H. J. (2019). Autistic adults anticipate and integrate meaning based on the speaker's voice: Evidence from eye-tracking and event-related potentials. Journal of Experimental Psychology: General, 149 (6), 1097–1115. https://doi.org/10.1037/xge0000705

Bastos, A. M., Martin Usrey, W., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. Neuron, 76(4), 695–711. https://doi.org/10.1016/j.neuron.2012.10.038

Braukmann, R., Ward, E., Hessels, R. S., Bekkering, H., Buitelaar, J. K., & Hunnius, S. (2018). Action prediction in 10-month-old infants at high and low familial risk for autism spectrum disorder. Research in Autism Spectrum Disorders, 49, 34–46. https://doi.org/10.1016/j.rasd.2018.02.004

Brisson, J., Warreyn, P., Serres, J., Foussier, S., & Adrien-Louis, J. (2012). Motor anticipation failure in infants with autism: a retrospective analysis of feeding situations. Autism, 16(4), 420–429. https://doi.org/10.1177/1362361311423385

Brock, J. (2012). Alternative Bayesian accounts of autistic perception: comment on Pellicano and Burr [Review of *Alternative Bayesian accounts of autistic perception: comment on Pellicano and Burr*]. Trends in Cognitive Sciences, 16(12), 573–574 author reply 574–575.

Brown, H., Friston, K., & Bestmann, S. (2011). Active inference, attention, and motor preparation. Frontiers in Psychology, 2, 218.

Cattaneo, L., Fabbri-Destro, M., Boria, S., Pieraccini, C., Monti, A., Cossu, G., & Rizzolatti, G. (2007). Impairment of actions chains in autism and its possible role in intention understanding. Proceedings of the National Academy of Sciences of the United States of America, 104(45), 17825–17830.

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.

Cheour, M., Alho, K., Ceponiené, R., Reinikainen, K., Sainio, K., Pohjavuori, M., … Näätänen, R. (1998). Maturation of mismatch negativity in infants. International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology, 29(2), 217–226.

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

Cox, D. J., Owens, J. M., Barnes, L., Moncrief, M., Boukhechba, M., Buckman, S., … Wotring, B. (2020). A pilot study comparing newly licensed drivers with and without autism and experienced drivers in simulated and on-road driving. Journal of Autism and Developmental Disorders, 50 (4), 1258–1268.

Dayan, P. (1993). Improving generalization for temporal difference learning: the successor representation. Neural Computation, 5 (4), 613–624. https://doi.org/10.1162/neco.1993.5.4.613

Deschrijver, E., Bardi, L., Wiersema, J. R., & Brass, M. (2016). Behavioral measures of implicit theory of mind in adults with high functioning autism. Cognitive Neuroscience, 7(1–4), 192–202.

Ego, C., Bonhomme, L., Orban de Xivry, J.-J., Da Fonseca, D., Lefèvre, P., Masson, G. S., & Deruelle, C. (2016). Behavioral characterization of prediction and internal models in adolescents with autistic spectrum disorders. Neuropsychologia, 91, 335–345.

Ewbank, M. P., Henson, R. N., Rowe, J. B., Stoyanova, R. S., & Calder, A. J. (2013). Different neural mechanisms within occipitotemporal cortex underlie repetition suppression across same and different-size faces. Cerebral Cortex, 23(5), 1073–1084.

Ewbank, M. P., Pell, P. J., Powell, T. E., von dem Hagen, E. A. H., Baron-Cohen, S., & Calder, A. J. (2017). Repetition suppression and memory for faces is reduced in adults with autism spectrum conditions. Cerebral Cortex, 27(1), 92–103.

Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. Frontiers in Human Neuroscience, 4, 215.

Font-Alaminos, M., Cornella, M., Costa-Faidella, J., Hervás, A., Leung, S., Rueda, I., & Escera, C. (2020). Increased subcortical

neural responses to repeating auditory stimulation in children with autism spectrum disorder. Biological Psychology, 149, 107807.

Foti, F., De Crescenzo, F., Vivanti, G., Menghini, D., & Vicari, S. (2015). Implicit learning in individuals with autism spectrum disorders: a meta-analysis. Psychological Medicine, 45(5), 897–910.

Friston, K. (2005). A theory of cortical responses. Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 360(1456), 815–836.

Friston, K. (2010). The free-energy principle: a unified brain theory? Nature Reviews Neuroscience, 11(2), 127–138.

Friston, K. (2018). Does predictive coding have a future? Nature Neuroscience, 21(8), 1019–1021. https://doi.org/10.1038/s41593-018-0200-7

Friston, K. J., Lawson, R., & Frith, C. D. (2013). On hyperpriors and hypopriors: comment on Pellicano and Burr [Review of *On hyperpriors and hypopriors*: *comment on Pellicano and Burr*]. Trends in Cognitive Sciences, 17(1), 1.

Fulceri, F., Tonacci, A., Lucaferro, A., Apicella, F., Narzisi, A., Vincenti, G., … Contaldo, A. (2018). Interpersonal motor coordination during joint actions in children with and without autism spectrum disorder: The role of motor information. Research in Developmental Disabilities, 80, 13–23.

Ganglmayer, K., Schuwerk, T., Sodian, B., & Paulus, M. (2020). Do children and adults with autism spectrum condition anticipate others' actions as goal-directed? A predictive coding perspective. Journal of Autism and Developmental Disorders, 50(6), 2077–2089.

Gardner, M. P. H., Schoenbaum, G., & Gershman, S. J. (2018). Rethinking dopamine as generalized prediction error. Proceedings of Biological Sciences/The Royal Society, 285, 1891. https://doi.org/10.1098/rspb.2018.1645

Gershman, S. J. (2018). The successor representation: Its computational logic and neural substrates. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 38 (33), 7193–7200.

Goldberg, M. C., Lasker, A. G., Zee, D. S., Garth, E., Tien, A., & Landa, R. J. (2002). Deficits in the initiation of eye movements in the absence of a visual target in adolescents with high functioning autism. Neuropsychologia, 40(12), 2039–2049. https://doi.org/10.1016/s0028-3932(02)00059-3

Gonzalez-Gadea, M. L., Chennu, S., Bekinschtein, T. A., Rattazzi, A., Beraudi, A., Tripicchio, P., … Ibanez, A. (2015). Predictive coding in autism spectrum disorder and attention deficit hyperactivity disorder. Journal of Neurophysiology, 114(5), 2625–2636.

Goris, J., Braem, S., Nijhof, A. D., Rigoni, D., Deschrijver, E., Van de Cruys, S., … 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.

Greene, R. K., Zheng, S., Kinard, J. L., Mosner, M. G., Wiesen, C. A., Kennedy, D. P., & Dichter, G. S. (2019). Social and nonsocial visual prediction errors in autism spectrum disorder. Autism Research: Official Journal of the International Society for Autism Research, 12(6), 878–883.

Grisoni, L., Moseley, R. L., Motlagh, S., Kandia, D., Sener, N., Pulvermüller, F., … Mohr, B. (2019). Prediction and mismatch negativity responses reflect impairments in action semantic processing in adults with autism spectrum disorders. Frontiers in Human Neuroscience, 13, 395.

Haker, H., Schneebeli, M., & Stephan, K. E. (2016). Can Bayesian theories of autism spectrum disorder help improve clinical practice? Frontiers in Psychiatry/Frontiers Research Foundation, 7, 107.

Hallett, V., Mueller, J., Breese, L., Hollett, M., Beresford, B., Irvine, A., … Simonoff, E. (2020). Introducing "predictive parenting": A feasibility study of a new group parenting intervention targeting emotional and behavioral difficulties in children with autism spectrum disorder. Journal of Autism and Developmental Disorders., 51, 323–333. https://doi.org/10.1007/s10803-020-04442-2

Happé, F., & Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. Journal of Autism and Developmental Disorders, 36(1), 5–25.

Herry, C., Bach, D. R., Esposito, F., Di Salle, F., Perrig, W. J., Scheffler, K., … Seifritz, E. (2007). Processing of temporal unpredictability in human and animal amygdala. Journal of Neuroscience, 27(22), 5958–5966. http://dx.doi.org/10.1523/jneurosci.5218-06.2007

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.

Karaminis, T., Cicchini, G. M., Neil, L., Cappagli, G., Aagten-Murphy, D., Burr, D., & Pellicano, E. (2016). Central tendency effects in time interval reproduction in autism. Scientific Reports, 6, 28570.

Kinard, J. L., Mosner, M. G., Greene, R. K., Addicott, M., Bizzell, J., Petty, C., … Dichter, G. S. (2020). Neural mechanisms of social and nonsocial reward prediction errors in adolescents with autism spectrum disorder. Autism Research: Official Journal of the International Society for Autism Research, 13(5), 715–728.

Kleinhans, N. M., Richards, T., Greenson, J., Dawson, G., & Aylward, E. (2016). altered dynamics of the FMRI response to faces in individuals with autism. Journal of Autism and Developmental Disorders, 46(1), 232–241.

Kolesnik, A., Begum Ali, J., Gliga, T., Guiraud, J., Charman, T., Johnson, M. H., … Team, B. A. S. I. S. (2019). Increased cortical reactivity to repeated tones at 8 months in infants with later ASD. Translational Psychiatry, 9(1), 46.

Kosaki, Y., & Watanabe, S. (2016). Impaired Pavlovian predictive learning between temporally phasic but not static events in autism-model strain mice. Neurobiology of Learning and Memory, 134, 304–316. https://doi.org/10.1016/j.nlm.2016.08.001

Krogh-Jespersen, S., Kaldy, Z., Valadez, A. G., Carter, A. S., & Woodward, A. L. (2018). Goal prediction in 2-year-old children with and without autism spectrum disorder: An eye-tracking study. Autism Research: Official Journal of the International Society for Autism Research, 11(6), 870–882.

Król, M., & Król, M. (2019). The world as we know it and the world as it is: Eye-movement patterns reveal decreased use of prior knowledge in individuals with autism. Autism Research, 12(9), 1386–1398. https://doi.org/10.1002/aur.2133

Kunchulia, M., Tatishvili, T., Lomidze, N., Parkosadze, K., & Thomaschke, R. (2017). Time-based event expectancies in

children with autism spectrum disorder. Experimental Brain Research: Experimentelle Hirnforschung. Experimentation Cerebrale, 235(9), 2877–2882.

Kunchulia, M., Tatishvili, T., Parkosadze, K., Lomidze, N., & Thomaschke, R. (2020). Children with autism spectrum disorder show increased sensitivity to time-based predictability. International Journal of Developmental Disabilities, 66, 1–8. https://doi.org/10.1080/20473869.2018.1564447

Landa, R. J., Haworth, J. L., & Nebel, M. B. (2016). Ready, set, go! Low anticipatory response during a dyadic task in infants at high familial risk for autism. Frontiers in Psychology, 7, 721.

Lawson, R. P., Mathys, C., & Rees, G. (2017). Adults with autism overestimate the volatility of the sensory environment. Nature Neuroscience, 20(9), 1293–1299.

Lawson, R. P., Rees, G., & Friston, K. J. (2014). An aberrant precision account of autism. Frontiers in Human Neuroscience, 8, 302.

Le Pelley, M. E. (2004). The role of associative history in models of associative learning: a selective review and a hybrid model. The Quarterly Journal of Experimental Psychology B, Comparative and Physiological Psychology, 57(3), 193–243.

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.

Lovaas, O. I., Koegel, R. L., & Schreibman, L. (1979). Stimulus overselectivity in autism: a review of research. Psychological Bulletin, 86(6), 1236–1254.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. Psychological Review, 82(4), 276–298. https://doi.org/10.1037/h0076778

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

Mayo, J., & Eigsti, I.-M. (2012). Brief report: a comparison of statistical learning in school-aged children with high functioning autism and typically developing peers. Journal of Autism and Developmental Disorders, 42(11), 2476–2485.

McDaniel, J. W. (1964). The effects of stimulus predictability upon auditory habituation and acquisition. [Doctoral dissertation, Texas Technical College]. Texas Tech University Libraries.

Millin, R., Kolodny, T., Flevaris, A. V., Kale, A. M., Schallmo, M.-P., Gerdts, J., … Murray, S. (2018). Reduced auditory cortical adaptation in autism spectrum disorder. eLife, 7, 7. https://doi.org/10.7554/eLife.36493

Mirza, M. B., Adams, R. A., Friston, K., & Parr, T. (2019). Introducing a Bayesian model of selective attention based on active inference. Scientific Reports, 9(1), 13915.

Mitchell, P., & Ropar, D. (2004). Visuo-spatial abilities in autism: A review. Infant and Child Development, 13(3), 185–198. https://doi.org/10.1002/icd.348

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2010). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. International Journal of Surgery, 8(5), 336–341.

Mosner, M. G., McLaurin, R. E., Kinard, J. L., Hakimi, S., Parelman, J., Shah, J. S., … Dichter, G. S. (2019). Neural mechanisms of reward prediction error in autism spectrum disorder. Autism Research and Treatment, 2019, 5469191.

Northrup, J. B. (2017). Contingency detection in a complex world. International Journal of Behavioral Development, 41 (6), 723–734. https://doi.org/10.1177/0165025416668582

Northrup, J. B., Libertus, K., & Iverson, J. M. (2017). Response to changing contingencies in infants at high and low risk for autism spectrum disorder. Autism Research: Official Journal of the International Society for Autism Research, 10(7), 1239–1248.

Obeid, R., Brooks, P. J., Powers, K. L., Gillespie-Lynch, K., & Lum, J. A. G. (2016). Statistical learning in specific language impairment and autism spectrum disorder: A meta-analysis. Frontiers in Psychology, 7, 1245.

Palumbo, L., Burnett, H. G., & Jellema, T. (2015). Atypical emotional anticipation in high-functioning autism. Molecular Autism, 6, 47.

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

Ploog, B. O. (2010). Stimulus overselectivity four decades later: a review of the literature and its implications for current research in autism spectrum disorder. Journal of Autism and Developmental Disorders, 40(11), 1332–1349.

Powell, P. S., Travers, B. G., Klinger, L. G., & Klinger, M. R. (2016). Difficulties with multi-sensory fear conditioning in individuals with autism spectrum disorder. Research in Autism Spectrum Disorders, 25, 137–146. https://doi.org/10.1016/j.rasd.2016.02.008

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nature Neuroscience, 2(1), 79–87. https://doi.org/10.1038/4580

Robertson, C. E., Kravitz, D. J., Freyberg, J., Baron-Cohen, S., & Baker, C. I. (2013). Tunnel vision: sharper gradient of spatial attention in autism. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 33(16), 6776–6781.

Ruiz-Martínez, F. J., Rodríguez-Martínez, E. I., Wilson, C. E., Yau, S., Saldaña, D., & Gómez, C. M. (2020). Impaired P1 habituation and mismatch negativity in children with autism spectrum disorder. Journal of Autism and Developmental Disorders, 50(2), 603–616.

Schallmo, M.-P., Kolodny, T., Kale, A. M., Millin, R., Flevaris, A. V., Edden, R. A. E., … Murray, S. O. (2020). Weaker neural suppression in autism. Nature Communications, 11(1), 2675.

Schubotz, R. I. (2016). Neural systems underlying the prediction of complex events. Cognitive Systems Monographs, 29, 81–92. https://doi.org/10.1007/978-3-319-22599-9_7

Schuwerk, T., Sodian, B., & Paulus, M. (2016). Cognitive mechanisms underlying action prediction in children and adults with autism spectrum condition. Journal of Autism and Developmental Disorders, 46(12), 3623–3639.

Scott-Van Zeeland, A. A., McNealy, K., Wang, A. T., Sigman, M., Bookheimer, S. Y., & Dapretto, M. (2010). No neural evidence of statistical learning during exposure to artificial languages in children with autism spectrum disorders. Biological Psychiatry, 68(4), 345–351.

Sheppard, E., van Loon, E., Underwood, G., & Ropar, D. (2016). Difficulties predicting time-to-arrival in individuals with autism spectrum disorders. Research in Autism Spectrum Disorders, 28, 17–23. https://doi.org/10.1016/j.rasd.2016.05.001

Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., … Held, R. M. (2014). Autism as a disorder of prediction. Proceedings of the National Academy of Sciences of the United States of America, 111(42), 15220–15225.

Sokolov, E. N. (1960). Neuronal models and the orienting reflex. In M. A. B. Brazier (Ed.), The central nervous system and behaviour, 3rd conference, (pp. 187–286). New York: Josiah Macy, Jr. Foundation.

Tam, F. I., King, J. A., Geisler, D., Korb, F. M., Sareng, J., Ritschel, F., … Ehrlich, S. (2017). Altered behavioral and amygdala habituation in high-functioning adults with autism spectrum disorder: an fMRI study. Scientific Reports, 7(1), 13611.

Teufel, C., & Fletcher, P. C. (2020). Forms of prediction in the nervous system. Nature Reviews Neuroscience, 21(4), 231–242.

Tewolde, F. G., Bishop, D. V. M., & Manning, C. (2018). Visual motion prediction and verbal false memory performance in autistic children. Autism Research: Official Journal of the International Society for Autism Research, 11(3), 509–518.

Thillay, A., Lemaire, M., Roux, S., Houy-Durand, E., Barthélémy, C., Knight, R. T., … Bonnet-Brilhault, F. (2016). Atypical brain mechanisms of prediction according to uncertainty in autism. Frontiers in Neuroscience, 10, 317.

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.

Turi, M., Karaminis, T., Pellicano, E., & Burr, D. (2016). No rapid audiovisual recalibration in adults on the autism spectrum. Scientific Reports, 6, 21756.

Turk-Browne, N. B., Scholl, B. J., & Chun, M. M. (2008). Babies and brains: habituation in infant cognition and functional neuroimaging. Frontiers in human neuroscience, 2, 16. https://doi.org/10.3389/neuro.09.016.2008

Utzerath, C., Schmits, I. C., Buitelaar, J., & de Lange, F. P. (2018). Adolescents with autism show typical fMRI repetition suppression, but atypical surprise response. Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 109, 25–34.

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: Official Journal of the International Society for Autism Research, 9(10), 1103–1113.

van Boxtel, J. J. A., & Lu, H. (2013). A predictive coding perspective on autism spectrum disorders. Frontiers in Psychology, 4, 19.

Van de Cruys, S., de Wit, L., Evers, K., Boets, B., & Wagemans, J. (2013). Weak priors versus overfitting of predictions in autism: Reply to Pellicano and Burr (TICS, 2012). I-Perception, 4(2), 95–97.

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

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.

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 Laarhoven, T., Stekelenburg, J. J., Eussen, M. L., & Vroomen, J. (2020). Atypical visual-auditory predictive coding in autism spectrum disorder: Electrophysiological evidence from stimulus omissions (1362361320926061). Autism: The International Journal of Research and Practice.

van Laarhoven, T., Stekelenburg, J. J., Eussen, M. L. J. M., & Vroomen, J. (2019). Electrophysiological alterations in motor-auditory predictive coding in autism spectrum disorder. Autism Research: Official Journal of the International Society for Autism Research, 12(4), 589–599.

Vivanti, G., Hocking, D. R., Fanning, P. A. J., Uljarevic, M., Postorino, V., Mazzone, L., & Dissanayake, C. (2018). Attention to novelty versus repetition: Contrasting habituation profiles in Autism and Williams syndrome. Developmental Cognitive Neuroscience, 29, 54–60.

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

Webb, S. J., Jones, E. J. H., Merkle, K., Namkung, J., Toth, K., Greenson, J., … Dawson, G. (2010). Toddlers with elevated autism symptoms show slowed habituation to faces. Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence, 16(3), 255–278.

Westerfield, M. A., Zinni, M., Vo, K., & Townsend, J. (2015). Tracking the sensory environment: an ERP study of probability and context updating in ASD. Journal of Autism and Developmental Disorders, 45(2), 600–611.

Wynn, K. (2008). Some innate foundations of social and moral cognition. The Innate Mind, 3, 330–347. https://doi.org/10.1093/acprof:oso/9780195332834.003.0017

Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. Neuron, 46(4), 681–692.

Zwart, F. S., Vissers, C. T. W. M., van der Meij, R., Kessels, R. P. C., & Maes, J. H. R. (2017). Autism: Too eager to learn? Event related potential findings of increased dependency on intentional learning in a serial reaction time task. Autism Research: Official Journal of the International Society for Autism Research, 10(9), 1533–1543.