Brain mechanisms of multisensory prediction and inference in neurotypical and autistic participants

Justin Chumbley

Translational neuromodeling Unit, ETH/UZH Zurich

1 Summary

Our brains solve a multitude of hard tasks posed by our environment, including perception, causal inference, decision-making and learning. It has been proposed that these diverse capabilities rest on a few common neuronal operations [4, 62, 32]. This possibility offers hope for a more unified, parsimonious explanation for atypical brain function [24, 45].

We propose a novel theoretical, and empirical examination of the common computations that may underlie these apparently diverse tasks, in both neurotypical and autistic spectrum disorder (ASD) participants. Our experiments will address fundamental open questions about the connections between multisensory perception, decision-making, causal inference & learning in the neurotypical brain. To ensure comparability across these different domains, all of our experimental paradigms are formally derived from a single learning task which requires subjects to identify useful patterns in simple multisensory sequences. Parallel testing of ASD participants will permit us to reduce any group differences to a common neuro-computational mechanism.

Our experiments draw in a simple way from the Bayesian brain hypothesis: the notion that neuronal networks solve diverse problems in a common manner, by encoding and updating beliefs about likely patterns in the environment, based on the available sensory data. We focus on ASD because faulty pattern recognition and sensory perception are core symptoms. Our experimental approach will combine established neuroimaging tools with novel, theoretically-inspired behavioral tasks. Our data analysis will feature simple, qualitative tests of our hypothesized neuro-computational operations, and their dysfunction in ASD. These tests will be complemented by quantitative tests, derived from more nuanced, explicitly neuro-computational and neuro-physiological models.

I have already piloted all the behavioral tasks and gathered and analyzed preliminary EEG pilots. I am based in the Translational Neuromodeling Unit, which operates at the intersection between medicine (UZH) and engineering (ETH) and runs its own research clinic, dedicated to studies of psychiatric patients. This work will benefit from collaboration with our resident psychiatrist, Helene Haker Rössler. Ethical approval for our studies will be fast and uncomplicated because they are not clinical trials.

2 State of research

2.1 Overview

A pattern is a relationship between the individual components of a whole. Patterns in our environment help us predict, understand and respond to the world. Real-world patterns are rarely perfect: they are obscured and corrupted in various ways. For example, on our second encounter with a stranger, he may have a new beard, obscuring some facial features. We may nonetheless recognize him, infer he has been too busy to shave and change our behavior accordingly. The ability to encode and reason about uncertain patterns in our environment is central to cognition: it underpins perception, causal inference, decision-making and learning.

We focus on the brain's ability to encode probable patterns because this is central to success across all domains of cognition. Multisensory perception requires us to predict likely patterns of sensation, experienced simultaneously at different sensors. Decision-making requires us to combine the predicted state of the world and the predicted value of that state: this configuration together determine the expected value of actions. Causal inference requires us to predict which pattern of causes explain an observed effect. As we will show, all these tasks can be modeled and solved by a common probabilistic pattern-representation system.

Perception [60], decision-making [26], causal inference [3] and learning [51] are altered in ASD – a neurodevelopmental cognitive disorder which prominently features irregular pattern processing. Notably, ASD sufferers perform differently from neurotypical subjects on tasks that involve pattern processing. For example, individuals with autism often outperform neurotypical individuals on tasks which require processing of local features, and may sometimes underperform on tasks which require global patterns [5, 50]. We will therefore contrast ASD sufferers with neurotypical individuals across all of these experiments with the aim of reducing group differences to a common pattern-recognition mechanism.

2.2 Multisensory learning & decision-making

An often-cited ecological advantage of multisensory neuronal codes is "change detection", i.e. enhanced perceptual sensitivity to surprising environmental changes [47, 36, 2, 29]. This enhancement arises because one can pool the unreliable information from each noisy sensory channel to gain more confidence about reality. Pooling is useful because environmental changes can produce correlated and therefore partially redundant signals in different sensory channels. This "congruence" between sensory signals in different channels is also emphasized in the literature on multisensory learning [56]. In a typical neurophysiology experiment, sensory channels are stimulated both alone and then in conjunction, in order to identify behaviors and neurons exhibiting multisensory facilitation. For example, visual and auditory speech cues may be presented individually or jointly, to assess any effect on speech comprehension. A typical observation is that weak or noisy unisensory stimuli can, when presented at the same time, pool superadditively to enhance detection.

Despite their fundamental importance to our understanding of perception, such paradigms do not provide a strong test for the role of multisensory processes in change detection. First, such detection tasks do not strictly require multisensory pooling. Pooling is not generally necessary for change detection, it simply helps when the individual signals are

weak, relative to noise. This is because in these tasks each individual sensory channel is sufficient for change detection, albeit with a lower success rate: mathematically, these tasks are "separable". Second, while there is important independent evidence of multisensory learning [1, 46, 28], these tasks do not explicitly examine the learning mechanisms by which regularities in the environment are internalized, so that any change is surprising. This would require measuring trial-by-trial neuronal activity as subjects learn to predict patterns in multisensory sequences, and as these patterns are subsequently violated by change. Third, because these detection tasks contain multisensory redundancy, they cannot be used to isolate brain responses to unexpected patterns. In other words, any measured surprise at an unexpected pattern is confounded with surprise at each of the components that constitute the pattern. This is particularly important for any attempt to unambiguously isolate surprise or "prediction-error" learning mechanisms implicated in pattern learning [33, 27]. Our experiments will address these shortfalls. In addition, our data on the role of multisensory surprise will bridge two different literatures on learning, which respectively emphasize the importance of multisensory learning techniques [56] and the importance of error or surprise [53, 32, 11].

The new DSM-5 diagnostic criteria for ASD includes sensory alterations. In fact, multisensory integration deficits have been hypothesized in ASD for decades and numerous therapies focus on facilitating sensory integration [54]. It is therefore remarkable how little data exist to assess this claim [59, 6]. There are important pointers that multisensory integration is compromised in ASD, such that sufferers have difficulty binding information coming from different sensory channels [55, 19, 30]. For example, autistic subjects do not perceptually bind audiovisual signals, with implications for language comprehension. No study to date has tested whether multisensory prediction or change detection are similarly compromised (e.g. as a downstream consequence of integration problems). This may be important because ASD sufferers typically show aberrant responses to environmental change: In particular, an "anxiously obsessive desire for sameness" [41]. It is currently not known whether this desire for sameness results from disturbing unpredictability in their multisensory experience.

2.3 Causal inference

Healthy subjects learn the patterns of causation that govern their environment [23, 43]. According to the most influential model of this causal learning, an internal association between cause and effect is updated whenever causal predictions are violated [57]. Unfortunately, roughly 40 different algorithmic models of causal learning exist, each explaining different aspects of the data [38]. Furthermore the majority are heuristic, i.e. they do not formally specify the computational goal of causal learning. Parsimonious computational principles may help unify and simplify models of causal inference [39, 35].

These theoretical principles have also greatly influenced recent models of perception [31, 52, 37, 42]. In such models, perception is seen as fundamentally a problem of causal inference, in which the computational goal is to infer the unobserved environmental causes of sensory stimulation. These models hold that humans represent the joint probability of causes and effects: causal inferences then result from simple operations on this joint representation. This joint probabilistic representation of both cause and effect is a defining characteristic of such "generative" models of cortical learning, and distinguishes them from other heuristics [25, 32]. This fundamental claim that the brain encodes joint patterns of

cause and effect in the service of causal inference is still an important open question [61], which we consider in our experiments below.

Altered causal reasoning and inference is central to one influential theory of autism [3]. By this theory, ASD sufferers fail to accurately infer the hidden beliefs and desires that cause others to act. Beyond social inference, there is a long-standing view that other aspects of everyday causal inference are compromised in ASD, but that these differences are not always easy to identify in the lab [34]. To date there are very few neuroimaging studies which contrast autistic and neurotypical causal inference [40]. Not one has directly examined the cortical computations underpinning multisensory causal inference in ASD (or healthy controls).

3 State of personal research

I am a statistician with a background in experimental psychology, economics and biology. I won a PhD scholarship and two Post-doctoral fellowships which gave me full freedom to identify and develop a research program in computational and behavioral neuroscience. My program includes theoretical and experimental work. First, I develop novel statistical tools for behavioral and big neuroimaging data [16, 15, 14, 18, 9, 8]. Second, my models and experiments examine learning and decision-making under uncertainty [12, 13, 17, 11], using tools from game theory, information theory, reinforcement and Bayesian learning, psychology and biology (electrophysiology, genetics, endocrinology). They ask whether people interact, search, explore, learn and generalize in an optimal manner, especially given conflicting observations, beliefs or interests. To date, my computational neuroscience projects involved developing mathematical models of prefrontal, striatal and hippocampal function [12, 13, 17, 11].

4 Detailed research plan

4.1 Experiments I: Pattern reversal paradigm (PRP)

Learning to detect recurrent or surprising patterns in our environment is crucial for survival. Remarkably, multisensory research has never unambiguously isolated this capability. We aim to understand how the brain learns to predict regular patterns, i.e. relationships between elemental features, and how it detects unexpected irregularities. We will do this by measuring how behavioral and brain responses to multisensory patterns change with learning. We propose a novel learning paradigm – a probabilistic pattern reversal paradigm (PRP) – in which the subject learns to predict the probability of upcoming multisensory patterns based on available cues. In this task the cue-conditional probability of each pattern periodically changes, simultaneously & without warning. Critically, the probability of each pattern **component** remains constant (Figure 1a). Thus, pattern reversal should selectively surprise any brain system which predicts multisensory patterns, because unpredicted patterns are now observed and predicted patterns are omitted. Pattern reversal will not however surprise any brain system which predicts the components of a pattern in isolation: due to our experimental design, such systems would be blind to reversal. By measuring neuronal surprise – or "prediction error" - we can isolate brain responses to unpredicted gestalt patterns that are not reducible or separable into their

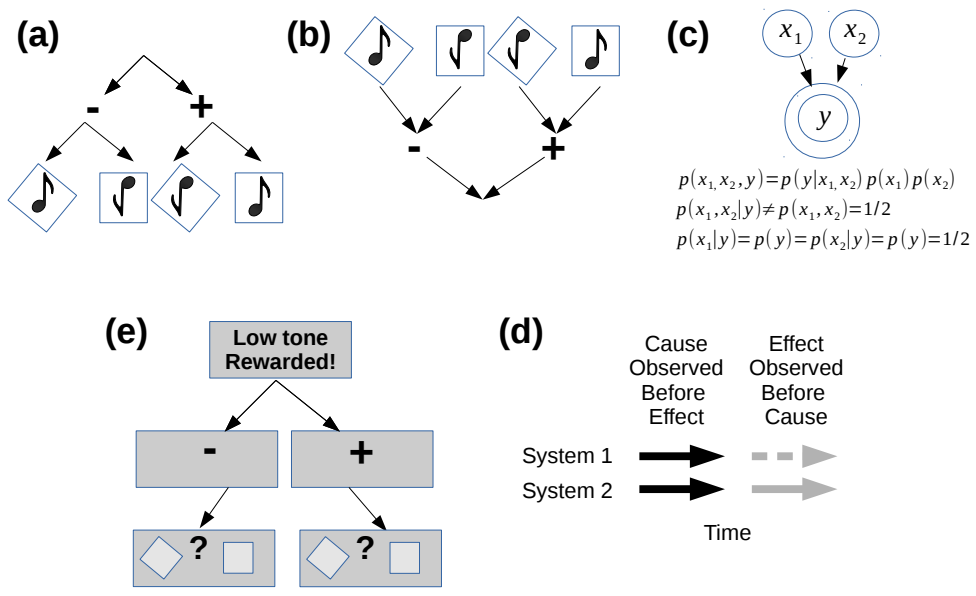

Figure 1: Our experimental paradigms. (a) Our probabilistic pattern reversal paradigm (PRP). On each trial we randomly sample a visual cue (+ or -) with equal probability, followed by a multisensory event. This multisensory event is composed of visual (square or diamond) and auditory (low or high tone) elements. In our design, none of these unisensory elements is ever particularly surprising because they always occur with equal probability regardless of cue reversal, i.e. it makes no difference if we secretly swap the "+" and "-" in this figure. Conditional on each cue however, a probabilistic pattern emerges: not all configurations are equally probable. In particular, diamond-low or square-high follow cue "-", and diamond-high and square-low follow cue "+". Having learned this contingency, subjects should be surprised if we reverse the contingencies because they now see unexpected patterns. (b) The opposite problem to (a). This contingency describes the multisensory "causes" of a visual "effect" (i.e. + or -). Reversing the contingencies, i.e. swapping outcomes "+" and "-", should lead to surprise, albeit unisensory surprise. (c) Tasks in (a) and (b) can be solved by learning separate, task-specific models: each solves their half of the problem. Alternatively, a single joint model can be learned. This joint model solves both problems and additionally permits new inferences. For example, having learned a forward or generative model by observing cause-effect relations in (b), it can later invert this model to infer the hidden cause of an observed effect as in (a). Using our causal inference reversal paradigm (CIRP), we test whether humans indeed do this. This joint model can be depicted generally as a causal directed acyclic graph. Here x_1 and x_2 are random variables indicating the multisensory cause, y is their effect. (d) In the CIRP, subjects are trained on forward causal models of System 1 and System 2. The sequence of observation is reversed in both systems. At the same time the causal contingencies in System 1 are also reversed. Causal inferences about System 1 should be violated relative to those of System 2, triggering measurable surprise. In a Bayes-optimal learner, this surprise should be formally identical to that following reversal in the PRP, providing comparability across experiments. We will measure EEG mismatch responses to this violation. (e) Our decision-making task asks subjects to complete patterns for reward.

elemental parts. We will measure this novel pattern surprise with EEG, in the knowledge that many other types of surprise or mismatch responses are readily observed with this methodology. By combining our passive learning paradigm with EEG, we can readily accommodate diverse and challenging populations, such as those encountered in adult and pediatric psychiatry.

4.1.1 Methods

Task: In our PRP, both multisensory patterns and their unisensory constituents are completely unpredictable without the cues. Cues do not predict the probability of any unisensory component in isolation. In contrast, each cue does predict the probability of certain multisensory patterns. In other words, having seen the cue, certain combinations or relationships of unisensory elements are more likely than others, despite the individual elements being equally likely.

Our design ensures that neurophysiological measures of pattern surprise cannot be attributed to habitation or the base presentation rate of any stimulus. The design ensures that novelty/familiarity of cues and outcomes are controlled because subjects see each cue and element the same number of times throughout the experiment. This is important because novelty might elicit confounding brain responses. Simple recency effects are constant because the presentation rate of each cue or element is the same.

Decision-making: Decision-making requires one to combine the predicted state of the world and the predicted value of that state: this configuration together determine the expected value of actions. We can therefore use decision-making as an index of effective pattern prediction in our task. To assess decision-making, we will use a simple choice task with real financial rewards (Figure 1d). After initial training in the PRP but before PRP reversal, subjects will decide how to complete patterns for reward. On each choice trial, subjects are first instructed about which sound will be rewarded. They then see a cue. Finally, they decide between shapes, see Figure 1d. If they choose the shape corresponding to the rewarded sound, then they win. Subjects who have successfully learned to predict patterns can win on every trial, otherwise they win at chance. We can therefore quantify how well subjects use pattern learning in the service of decision-making, using their attained reward as a behavioral index of learning.

Brain ERP: We will quantify EEG responses [7, 33] to surprising versus non-surprising patterns (before versus after reversal).

Brain connectivity and plasticity: We will use our EEG data to infer how brain networks learn to predict multisensory features in our PRP. To do this, we draw on established data-analytic tools for EEG time-series: so called "dynamic causal models" of brain connectivity [49]. These tools permit us to model activity in auditory regions, visual regions and "cross-modal" regions, and their dynamic interactions. Critically these tools permit us to identify the physiological cause of surprise-dependent connectivity changes in brain, and relate it to prediction error theories of learning. This gives us a direct handle on the mechanisms of adaptive pattern prediction. We will use fine-scale computational models of learning dynamics to dissect the algorithms which might implement multisensory learning, relating them to important prior work [44, 62, 4, 42].

Participants: We will collect data from 150 neurotypical and 150 ASD subjects. This permits us to publish a paper both on the basic mechanisms, as well as any observed aberration in ASD. To help with data collection, we will recruit a EEG research techni-

cian for 2 years. Our sample size is based on the following power calculation: To detect a 3 microampere ERP group difference, assuming a within-group standard deviation of 4 (constant across groups), a sample size of 75 participants per study per group should give us power of 99.8%. Inclusion criterea: $\geq$18 years, written informed consent, in-house clinical diagnosis of Autism spectrum disorder. Here in switzerland, patients are diagnosed according to ICD-10. This means, they mostly have a diagnosis of Asperger Syndrome (or atypical autism). Exclusion criterea: acute psychiatric disorders (i.e. the most prominent ASD comorbidities are substance abuse, depressive episode, psychotic episode), antipsychotic drugs or stimulants (e.g. dopamine antagonists or methylphenidate) within the last 48h (dopamine may be involved in our learning task) and any acute medical condition. We will take other measures - including the temporal binding window - in both ASD and neurotypical, subjects to better understand group differences and relate them to influential probabilistic accounts of ASD [10, 48, 50].

4.2 Experiments II: Causal inference reversal paradigm (CIRP)

In our PRP paradigm above, subjects learned cue-conditional patterns: a multisensory pattern only existed conditional on the cue. It is remarkable that this same capacity to encode conditional patterns is central to recent theories of causal inference. In these theories, subjects internally model causal relations in the world and infer unobserved causes or effects. For example, whenever a single observed effect can arise from two hidden causes, subjects are uncertain about which cause was responsible, until they observe the cause itself. This (inverse) relationship between the plausibility of two competing causal explanations can invoke exactly the same probabilistic patterns of anticipation as seen in the PRP above, where the "cue" is simply the observed effect to be explained. We aim to find evidence for this form of pattern processing, which is a signature of Bayesian causal inference.

Our causal inference reversal paradigm (CIRP) uses two facts to achieve this. First, reversing causal contingencies should generate surprise at previously unexpected events. Second, in principle a subject may observe cause and effect in any order: the sequence of causation need not be the sequence of observation. As explained next, our experimental design uses both of these facts, jointly reversing both the order of observation and the contingency.

4.2.1 Methods

Task: On each trial of the CIRP, subjects will first observe a multisensory "cause", followed by it's unisensory "effect", in two - formally identical - causal systems (Figure 1b and Figure 1d). The causal contingency in both systems is the exact inverse of that in the PRP (Figure 1a), a fact which will permit comparability across the two experiments. After interleaved training on both causal systems, we will reverse the sequence in which subjects observe cause and effect, presenting the effect first (Figure 1e). This permits subjects to infer the cause of this observed effect, before actually observing it. We will measure EEG surprise responses to violations of multisensory causal inference, as discussed next. To do this, in one of the two systems, concurrently with this sequence-reversal, we will also reverse the causal-contingencies. This contingency-reversal will surprise any brain center which had correctly inferred likely multisensory cause of the observed effect (by

inverting a learned, forward or generative model). Thus, any differential brain response to contingency reversal must reflect surprising causal inferences.

Other dependent measures: We will use an analogous decision task to the PRP, and analogous EEG acquisition and analysis (see above).

5 Schedule and milestones

We anticipate 5-6 papers from this work. The intended schedule for years 1 & 2 will be as follows. Months 1-2: Programming, Month 3-5: Behavioral and EEG (Healthy controls), Month 6: Data Analysis, Months 7-8: Preparation of manuscript, Month 8: Preparation of ASD EEG studies, Months 9-10: ASD + neurotypical EEG, Months 11-12: Analysis and preparation of manuscript for publication, based on the pipeline established from the preceding study with control subjects. In the final year, we will not hire a technician. I will gather any outstanding EEG data. In this year I will also write theoretical papers that synthesize and model all of our findings, putting them in the context of existing theories of neuronal computation.

6 Relevance and impact

Basic science: The goal of computational neuroscience is to meaningfully parse the apparent diversity and complexity of brain structure and function. Following this ideal, our simple experiments are original in attempting to follow a unifying computational theme across multiple different domains. This approach promises to open new links between traditionally distinct academic fields. For example, the "multisensory integration" literature does not address the learning algorithms which the brain uses to update multisensory predictions. Conversely, the canonical models from learning theory do not address how the brain learns to predict multisensory patterns. Our work adds to both fields and helps bridge them. **Applied science:** Multisensory function is disrupted in many psychiatric populations. In autism and schizophrenia, for example, surprising multisensory events are at the core of theories of aberrant perception [21, 20, 22, 58]. Our tasks offer a new and simple test of this theory, which is easy to administer in challenging populations. **Clinical:** Early intervention in ASD could improve development. Current diagnostic methods rest partly on verbal clinical assessment. This restricts the age at which autism can be diagnosed due to language barriers. There is a need for a reliable clinical tools for the diagnosis of autism in young children. EEG is low cost, non-invasive, and suitable for clinical applications, as well as academic ones such as ours.

7 Bibliography

References

[1] Wendy J Adams, Erich W Graf, and Marc O Ernst. Experience can change the'light-from-above'prior. *Nature neuroscience*, 7(10):1057–1058, 2004.

[2] Dora E Angelaki and Kathleen E Cullen. Vestibular system: the many facets of a multimodal sense. *Annu. Rev. Neurosci.*, 31:125–150, 2008.

[3] Simon Baron-Cohen, Alan M Leslie, and Uta Frith. Does the autistic child have a 'theory of mind'? *Cognition*, 21(1):37–46, 1985.

[4] Jeffrey M Beck, Peter E Latham, and Alexandre Pouget. Marginalization in neural circuits with divisive normalization. *The Journal of Neuroscience*, 31(43):15310–15319, 2011.

[5] Marlene Behrmann, Galia Avidan, Grace Lee Leonard, Rutie Kimchi, Beatriz Luna, Kate Humphreys, and Nancy Minshew. Configural processing in autism and its relationship to face processing. *Neuropsychologia*, 44(1):110–129, 2006.

[6] Alice B Brandwein, John J Foxe, John S Butler, Hans-Peter Frey, Juliana C Bates, Lisa H Shulman, and Sophie Molholm. Neurophysiological indices of atypical auditory processing and multisensory integration are associated with symptom severity in autism. *Journal of autism and developmental disorders*, 45(1):230–244, 2015.

[7] Alice B Brandwein, John J Foxe, John S Butler, Natalie N Russo, Ted S Altschuler, Hilary Gomes, and Sophie Molholm. The development of multisensory integration in high-functioning autism: high-density electrical mapping and psychophysical measures reveal impairments in the processing of audiovisual inputs. *Cerebral Cortex*, page bhs109, 2012.

[8] K.H. Brodersen, J. Daunizeau, C. Mathys, J.R. Chumbley, J.M. Buhmann, and K.E. Stephan. Variational bayesian mixed-effects inference for classification studies. *NeuroImage*, 2013.

[9] K.H. Brodersen, C. Mathys, J.R. Chumbley, J. Daunizeau, C.S. Ong, J.M. Buhmann, and K.E. Stephan. Bayesian mixed-effects inference on classification performance in hierarchical datasets. *Journal of Machine Learning Research*, 2012.

[10] Jason S Chan and Marcus J Naumer. Explaining autism spectrum disorders: central coherence vs. predictive coding theories. *Journal of neurophysiology*, 112(11):2669–2671, 2014.

[11] J.R. Chumbley, C.J. Burke, K.E. Stephan, K.J. Friston, P.N. Tobler, and E Fehr. Surprise beyond prediction error. *Human brain mapping*, 2014.

[12] J.R. Chumbley, R.J. Dolan, and K.J. Friston. Attractor models of working memory and their modulation by reward. *Biological Cybernetics*, 98(1):11–18, 2008.

[13] J.R. Chumbley, G. Flandin, D.R. Bach, J. Daunizeau, E. Fehr, RJ Dolan, and KJ Friston. Learning and generalization under ambiguity: An fmri study. *PLoS Computational Biology*, 8(1), 2012.

[14] J.R. Chumbley, G. Flandin, ML Seghier, and KJ Friston. Multinomial inference on distributed responses in spm. *NeuroImage*, 53(1):161–170, 2010.

[15] J.R. Chumbley and K.J. Friston. False discovery rate revisited: Fdr and topological inference using gaussian random fields. *NeuroImage*, 44(1):62–70, 2009.

[16] J.R. Chumbley, K.J. Friston, T. Fearn, and S.J. Kiebel. A metropolis-hastings algorithm for dynamic causal models. *NeuroImage*, 38(3):478–487, 2007.

[17] J.R. Chumbley, PN Tobler, and E Fehr. Fatal attraction: Ventral striatum predicts costly choice errors in humans. *NeuroImage*, 2013.

[18] J.R. Chumbley, K. Worsley, G. Flandin, and K. Friston. Topological fdr for neuroimaging. *NeuroImage*, 49(4):3057–3064, 2010.

[19] Olivier Collignon, Genevieve Charbonneau, Frederic Peters, Marouane Nassim, Maryse Lassonde, Franco Lepore, Laurent Mottron, and Armando Bertone. Reduced multisensory facilitation in persons with autism. *cortex*, 49(6):1704–1710, 2013.

[20] Philip R Corlett, Graham K Murray, Garry D Honey, Michael RF Aitken, David R Shanks, Trevor W Robbins, Edward T Bullmore, Anthony Dickinson, and Paul C Fletcher. Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. *Brain*, 130(9):2387–2400, 2007.

[21] PR Corlett, GK Murray, GD Honey, MRF Aitken, DR Shanks, TW Robbins, ET Bullmore, A Dickinson, and PC Fletcher. Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. *Brain*, 130(9):2387–2400, 2007.

[22] PR Corlett, JR Taylor, XJ Wang, PC Fletcher, and JH Krystal. Toward a neurobiology of delusions. *Progress in Neurobiology*, pages 345–369, 2010.

[23] David Danks, Thomas L Griffiths, and Joshua B Tenenbaum. Dynamical causal learning. In *Advances in neural information processing systems*, pages 67–74, 2002.

[24] Greg Davis and Kate Plaisted-Grant. Low endogenous neural noise in autism. *Autism*, page 1362361314552198, 2014.

[25] Peter Dayan and Laurence F Abbott. *Theoretical neuroscience*. Cambridge, MA: MIT Press, 2001.

[26] Benedetto De Martino, Neil A Harrison, Steven Knafo, Geoff Bird, and Raymond J Dolan. Explaining enhanced logical consistency during decision making in autism. *The Journal of Neuroscience*, 28(42):10746–10750, 2008.

[27] Martin J Dietz, Karl J Friston, Jason B Mattingley, Andreas Roepstorff, and Marta I Garrido. Effective connectivity reveals right-hemisphere dominance in audiospatial perception: implications for models of spatial neglect. *The Journal of Neuroscience*, 34(14):5003–5011, 2014.

[28] Mark O Ernst, Christoph Lange, and Fiona N Newell. Multisensory recognition of actively explored objects. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 61(3):242, 2007.

[29] Christopher R Fetsch, Alexandre Pouget, Gregory C DeAngelis, and Dora E Angelaki. Neural correlates of reliability-based cue weighting during multisensory integration. *Nature neuroscience*, 15(1):146–154, 2012.

[30] Jennifer H Foss-Feig, Leslie D Kwakye, Carissa J Cascio, Courtney P Burnette, Haleh Kadivar, Wendy L Stone, and Mark T Wallace. An extended multisensory temporal binding window in autism spectrum disorders. *Experimental Brain Research*, 203(2):381–389, 2010.

[31] K. Friston, J. Kilner, and L. Harrison. A free energy principle for the brain. *Journal of Physiology-Paris*, 100(1-3):70–87, 2006.

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

[33] Marta I Garrido, James M Kilner, Klaas E Stephan, and Karl J Friston. The mismatch negativity: a review of underlying mechanisms. *Clinical neurophysiology*, 120(3):453–463, 2009.

[34] Adam E Green, Lauren Kenworthy, Maya G Mosner, Natalie M Gallagher, Edward W Fearon, Carlos D Balhana, and Benjamin E Yerys. Abstract analogical reasoning in high-functioning children with autism spectrum disorders. *Autism Research*, 7(6):677–686, 2014.

[35] Thomas L Griffiths and Joshua B Tenenbaum. Theory-based causal induction. *Psychological review*, 116(4):661, 2009.

[36] Yong Gu, Dora E Angelaki, and Gregory C DeAngelis. Neural correlates of multisensory cue integration in macaque mstd. *Nature neuroscience*, 11(10):1201–1210, 2008.

[37] L. M. Harrison, A. Duggins, and K. J. Friston. Encoding uncertainty in the hippocampus. *Neural Networks*, 19(5):535–546, 2006. Times Cited: 19.

[38] Masasi Hattori and Mike Oaksford. Adaptive non-interventional heuristics for covariation detection in causal induction: Model comparison and rational analysis. *Cognitive Science*, 31(5):765–814, 2007.

[39] Keith J Holyoak and Patricia W Cheng. Causal learning and inference as a rational process: The new synthesis. *Annual review of psychology*, 62:135–163, 2011.

[40] Rajesh K Kana, Lauren E Libero, Christi P Hu, Hrishikesh D Deshpande, and Jeffrey S Colburn. Functional brain networks and white matter underlying theory-of-mind in autism. *Social cognitive and affective neuroscience*, 9(1):98–105, 2014.

[41] Leo Kanner. *Autistic disturbances of affective contact*. New York, 1943.

[42] Konrad P Körding, Ulrik Beierholm, Wei Ji Ma, Steven Quartz, Joshua B Tenenbaum, and Ladan Shams. Causal inference in multisensory perception. *PLoS one*, 2(9):e943, 2007.

[43] Hongjing Lu, Randall R Rojas, Tom Beckers, and Alan Yuille. Sequential causal learning in humans and rats. *Department of Statistics, UCLA*, 2011.

[44] Paul R MacNeilage, Narayan Ganesan, and Dora E Angelaki. Computational approaches to spatial orientation: from transfer functions to dynamic bayesian inference. *Journal of neurophysiology*, 100(6):2981–2996, 2008.

[45] Laurent Mottron. The costs of generality. *Autism*, page 1362361314558280, 2014.

[46] Saskia K Nagel, Christine Carl, Tobias Kringe, Robert Martin, and Peter Koenig. Beyond sensory substitution: learning the sixth sense. *Journal of neural engineering*, 2(4):R13, 2005.

[47] Tomokazu Ohshiro, Dora E Angelaki, and Gregory C DeAngelis. A normalization model of multisensory integration. *Nature neuroscience*, 14(6):775–782, 2011.

[48] Elizabeth Pellicano and David Burr. When the world becomes too real: a bayesian explanation of autistic perception. *Trends in cognitive sciences*, 16(10):504–510, 2012.

[49] William D Penny, Karl J Friston, John T Ashburner, Stefan J Kiebel, and Thomas E Nichols. *Statistical parametric mapping: the analysis of functional brain images: the analysis of functional brain images*. Academic press, 2011.

[50] Kate Plaisted, Lisa Saksida, Jose Alcantara, and Emma Weisblatt. Towards an understanding of the mechanisms of weak central coherence effects: Experiments in visual configural learning and auditory perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 358(1430):375–386, 2003.

[51] Kate C Plaisted. Reduced generalization in autism: An alternative to weak central coherence. 2001.

[52] R. P. N. Rao and D. H. Ballard. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1):79–87, 1999.

[53] R Rescorla and A Wagner. Variations in the effectiveness of reinforcement and nonreinforcement. *New York: Classical Conditioning II: Current Research and Theory, Appleton-Century-Crofts*, pages 113–123, 1972.

[54] Sylvia Rodger and Helene J Polatajko. Occupational therapy for children with autism. In *Comprehensive Guide to Autism*, pages 2297–2314. Springer, 2014.

[55] Natalie Russo, John J Foxe, Alice B Brandwein, Ted Altschuler, Hilary Gomes, and Sophie Molholm. Multisensory processing in children with autism: high-density electrical mapping of auditory–somatosensory integration. *Autism Research*, 3(5):253–267, 2010.

[56] Ladan Shams and Aaron R Seitz. Benefits of multisensory learning. *Trends in cognitive sciences*, 12(11):411–417, 2008.

[57] David R Shanks, Susan M Pearson, and Anthony Dickinson. Temporal contiguity and the judgement of causality by human subjects. *The Quarterly Journal of Experimental Psychology*, 41(2):139–159, 1989.

[58] Sukhwinder S Shergill, Gabrielle Samson, Paul M Bays, Chris D Frith, and Daniel M Wolpert. Evidence for sensory prediction deficits in schizophrenia. 2014.

[59] Elizabeth G Smith and Loisa Bennetto. Audiovisual speech integration and lipreading in autism. *Journal of Child Psychology and Psychiatry*, 48(8):813–821, 2007.

[60] Ryan A Stevenson, Justin K Siemann, Brittany C Schneider, Haley E Eberly, Tiffany G Woynaroski, Stephen M Camarata, and Mark T Wallace. Multisensory temporal integration in autism spectrum disorders. *The Journal of Neuroscience*, 34(3):691–697, 2014.

[61] Christopher Summerfield and Etienne Koechlin. A neural representation of prior information during perceptual inference. *Neuron*, 59(2):336–347, 2008.

[62] Nienke van Atteveldt, Micah M Murray, Gregor Thut, and Charles E Schroeder. Multisensory integration: flexible use of general operations. *Neuron*, 81(6):1240–1253, 2014.