**Metacognitive and noradrenergic alterations in autistic adults.**

**Ainslie Johnstone1, Karl Friston2, Geraint Rees\*,2,3, & Rebecca P. Lawson\*,†,4.**

1 Department for Clinical and Movement Neuroscience, UCL Queen Square Institute of Neurology, University College London, London, WC1N 3BG.

2 Wellcome Centre for Human Neuroimaging, University College London, 12 Queen Square, London WC1N 3AR

3 Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London, WC1N 3AZ.

4 Department of Psychology, University of Cambridge, Downing Street, Cambridge, CB2 3EB.

Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London, WC1N 3AR.

\*denotes shared authorship

† corresponding author

**Abstract**

People with autism spectrum disorder (ASD) experience the world differently, but the neurocognitive mechanisms that underlie these differences are poorly understood. In neurotypical (NT) adults, everyday perceptions are accompanied a sense of reliability, or confidence, which promotes flexible learning and adaptive behaviour. One novel hypothesis is that perception is not compromised in ASD *per se,* but rather the ability to identify the fidelity of one’s own perception; i.e. metacognition. To test this, we measured perceptual metacognitive efficiency in adults with and without ASD using a visual discrimination task in combination with hierarchical decision theoretic modelling. Our results reveal that adults with ASD exhibit reduced correspondence between accuracy and confidence in perception, despite equivalent task performance. We then demonstrate that the relationship between confidence and pupil dilation (a proxy for noradrenergic function) shows reduced dynamic range in autism; diminished response pre-perceptually, and enhanced response post-perceptually. Our results offer empirical support for previously hypothesised, but never tested, metacognitive and noradrenergic alterations in ASD, offering mechanistic insights into the non-social features of the condition.

**Key words**: Autism; Metacognition; Perception; Bayesian; Precision: Noradrenaline.

**Introduction**

Autism Spectrum Disorder (ASD[[1]](#endnote-1)) is a developmental condition with a complex genetic basis 1, that is characterised by differences in social interaction, restricted interests, repetitive behaviours, and sensory sensitivities 2. There is no account of ASD that spans cognitive, biological, and behavioural levels of description, but recent theories that draw on Bayesian models of brain function hold great promise 3–7. Some of these accounts focus on the brains ascending neuromodulatory pathways and the cognitive processes that depend upon their integrity, such as meta-learning 5,7. Specifically, these accounts emphasise differences in how fluctuations in confidence, or uncertainty, produce alterations to the cortical gain on sensory inputs, to give rise to the sensory and social features of the condition 5,7.

Confidence helps us determine how much weight we should give to different sources of information in the sensory environment, which enables us to update our predictions and adjust our actions. In neurotypical (NT) adults everyday perceptual judgements are accompanied by an instinctive and effortless sense of confidence, which is known to promote flexible learning and adaptive behaviour in perceptual decision making 8. It is possible to quantify the quality of someone’s confidence judgements, i.e. metacognitive ability, by measuring their fidelity with respect to objective task performance 9. Someone with good metacognitive ability is able to accurately identify and report on fluctuations in their performance. Metacognition shows substantial individual differences even when task performance is matched, and dimensional investigations of metacognitive ability in relation to trait psychopathology underwrite the fact that such individual differences have real-world consequences for mood and behaviour 10.

We previously hypothesised metacognitive difficulties in ASD 4, but this has yet to be formally tested under a decision theoretic framework that can rule out systematic response biases which may mimic metacognitive impairments 9. We recently demonstrated that the ability to learn predictive relationships under conditions of uncertainty is compromised in ASD, which affects the ability to build precise prior expectations 11. Metacognitive ability is highly predictive of one’s ability to learn probabilistic relationships 12, and in tasks that explicitly model learning, trial-by-trial confidence ratings have been shown to act as a weighting factor, much like a learning rate 13. We therefore reasoned in a previous hypothesis article, that in ASD the quality of confidence judgments (that is, the ability to distinguish between correct responses and errors) might be altered 4. Here we test whether perceptual metacognitive efficiency is diminished in adults with ASD. We hypothesised that when objective perceptual performance is matched between the groups, individuals with ASD would show poorer insight into their perceptual ability.

In terms of neurobiology, the ascending neuromodulator pathways are well placed to enact rapid changes in neural gain in the face of changing confidence 8,14. In particular noradrenaline is known to change the gain or signal to noise ratio in sensory cortex 15 and has increasingly shown to play an important role in metacognition. Confidence ratings in humans correlate with pupil dilation 16,17, a proxy for noradrenergic function 18, and blocking the action of noradrenaline changes both learning in uncertain environments 19, and also improves metacognition 20. We previously made the prediction that*,* at a physiological level, autistic metacognitive differences might therefore be related to neuromodulatory control 4. Here additionally assess whether adults with ASD show difference in confidence-linked metrics of noradrenergic function.

**Materials & Methods**

*Participants*

We tested adults with ASD (n=24) and NT adults (n=21) matched for age, gender and IQ (Table 1). See Supplementary Materials and Table S1 for full details.

*Stimuli and Procedure*

The task, previously used to assess metacognition in clinical groups 21,22, consisted of two-alternative forced choice judgements about which of two briefly presented circles contained more dots, followed by confidence ratings about that decision on each trial (Fig. 1). On each trial one of the circles, selected randomly, contained 50 dots and the other contained a number of dots bounded between 1 and 100. The difference in dot number between the two circles (Δ dots) was titrated to maintain constant performance for each participant (~75% accuracy) using a one-up two-down staircase as employed previously 21,22. After two consecutive correct responses Δ dots was decreased by one dot and after one incorrect response Δ dots was increased by one dot. This staircase equated the difficulty of the task between individuals. See Supplementary Materials for full details.

*Eye tracking*

As a measure of neuromodulatory responses, pupil size was measured with an infrared eye tracker (SR-Research Eyelink 1000) tracking at 1000 Hz. Chin and forehead were stabilised using a table mounted head rest.

*Statistical Power*

We powered this study to detect a difference between two independent means (2-tailed) for the effect size obtained in 22 (Cohen’s d = 1.397) at an α of 0.05 using G-Power. This calculation indicated a minimum sample size of 15 participants per group for a power of 95%.

*Data Analysis*

For full details of our data analysis procedures see the Supplementary Materials.

Performance was quantified as the proportion correct responses and the difficulty level (e.g. Δ dots threshold) was calculated as the mean number of dots added or subtracted to the target stimulus by the staircase procedure to maintain 75% performance.

To estimate metacognitive abilities we computed meta-d’ 23. Meta-d’ was fit to each participant’s data using MATLAB code freely available at: <http://www.columbia.edu/~bsm2105/type2sdt/>. Metacognitive efficiency was assessed as the ratio of meta-d’/d’ 9,21,22.

In a complementary analysis we estimated meta-d’/d’ using a hierarchical Bayesian version of the standard metacognitive efficiency model (HMeta-d toolbox 24, <https://github.com/smfleming/HMM>) which allows estimation and comparison of *group-level* parameters. The parameters were estimated using Markov-Chain Monte-Carlo methods (MCMC, here: 3 chains of 10’000 samples each, burn-in of 1000 samples) as implemented in JAGS (<http://mcmc-jags.sourceforge.net>).

All statistical analyses of eye-tracking data were performed in MATLAB (Mathworks). For the pupillometry analyses we applied a general linear modelling approach to estimate the encoding of confidence in pupil size across time (while controlling for confounding variables).

**Results**

A two-up-one-down staircase successfully equated task performance between the groups (t (43) = 0.36, P = 0.72; Fig. 2A) and there was no difference in the perceptual difficulty (average Δ dots) necessary to maintain equal performance between the groups (t (43) = 0.15, P = 0.88; Fig. 2B). This equivalent perceptual performance was further confirmed by d’ scores (type 1 perceptual sensitivity; t (43) = -0.31, P = 0.76; Fig. 2C), and there was no significant difference in the average level of confidence reported (t (43) = 0.04, P = 0.96; Fig. 2D) suggesting that all participants understood how to use the confidence scale.

We used a signal detection theoretic modelling approach to assess metacognitive ability (Supplementary Materials). Despite equivalent perceptual performance, metacognitive sensitivity (meta-d’; t (43) = -2.29, P=0.027; Fig. 3A) and metacognitive efficiency (meta-d’/d’; t (43) = -2.13, P = 0.039) were significantly lower in the ASD group relative to the NT group (Fig. 3B). This result was confirmed on log transformed meta-d’/d’ (t (43) = -2.20, P = 0.035) and bootstrap t-test (P = 0.043 [ 95% CI -0.38 - -0.02]). Since the ratio of meta-d’/d’ quantifies metacognitive accuracy in units of task performance (d’) it is theoretically meaningful to compare this quantity against the optimal meta-d’/d’ score of 1. The NT group’s metacognitive efficiency score was 0.88, e.g. 88% of optimal, but did not statistically differ from optimal (t (20) = -1.79, P = 0.09; bootstrap P=0.098 [95% CI -0.25 – 0.02]). The ASD group, however, had a lower metacognitive efficiency score of 0.67, e.g. 67% of optimal, that did significantly differ from optimal (t (23) = -4.59, P<0.001; bootstrap t-test P=0.001 [95% CI -0.44 – -0.20]). A complementary analysis using a recently developed hierarchical method to fit group-level parameters for meta-d’/d (Supplementary Materials), confirmed this core group difference in metacognitive efficiency (Fig. 3C&D; P (group difference meta-d’/d’) < 0 = 0.0067).

There was no group difference in basic pupil response to visual stimuli (Supplemental Figure S1). In NT participants we replicated previous findings that confidence is encoded in pupil size in a biphasic manner, with opposite effects around the time of the visual stimulus and the confidence judgement 17. In NT participants, stimuli that will subsequently be rated with high confidence are preceded by an *increase* in pupil size before the stimulus appears, indicating that neuromodulatory or arousal state affects the ‘readiness’ to process sensory inputs which influences subsequent confidence. This effect was diminished in ASD (Fig. 4A). Conversely, high confidence ratings elicit a sustained *decrease* in pupil size post-perceptually, after the confidence rating itself. This latter effect is consistent with the role of noradrenaline in signalling confidence in prior perceptual decisions 14, i.e. when uncertainty in what you’ve just seen is low, confidence ratings are high, and the action of noradrenaline (pupil size) is suppressed. Accordingly, high confidence resulted in a decrease in pupil size in both groups; however, the latency of this effect was diminished in the ASD group with pupil size quickly increasing to baseline (Fig. 4B). This indicates elevated post-perceptual encoding of confidence in pupil size, consistent with recent work linking larger pupil sizes to precision-weighted learning dynamics in ASD 11. Taken together, the effects of confidence on pupil dilation at both time-points in the trial indicate reduced dynamic range in autism, evidence for diminished neuromodulatory control.

**Discussion**

We found that adults with autism exhibit reduced metacognitive efficiency during perceptual decision making, relative to typically developed adults. The ASD group showed otherwise equivalent perceptual functioning, with comparable overall task performance, overall task confidence and overall task difficulty to the NT group. Our behavioural results identify a specific difficulty with the trial-by-trial match between confidence and perceptual performance in adults with autism, consistent with poorer insight into confidence fluctuations in ASD. Under the Predictive Coding framework, confidence reflects the precision (or inverse variance) of a higher order belief about sensory states 25. Precision is regulated by the brains ascending neuromodulatory pathways, to nuance the gain on cortical responses 26. Predictive Coding accounts of autism posit differential precision dynamics 5,7, which we previously suggested would manifest as alterations in confidence-linked noradrenergic function in autism 4. Consistent with this prediction our pupillometry results demonstrate altered neuromodulatory control autism, an attenuation of both pre- and post- stimulus effects of confidence on pupil size.

In explaining sensory perception in ASD, prominent cognitive theories have focussed on imbalances between local and global sensory processing; most obviously the weak central coherence hypothesis (WCC; Happé and Frith, 2006), and the Enhanced Perceptual Functioning theory (EPF; Mottron *et al.*, 2006). Both theories fall short in explaining the results of the present study. With regard to the EPF theory, we have shown not only that perceptual performance is the same between the groups (e.g. successful staircasing of performance; Fig 2A), but crucially that the task difficulty required to maintain equivalent performance between the groups was also exactly the same (Fig. 2B). An account of perceptual differences in ASD that emphasised enhanced bottom-up processing would hypothesise that in the ASD group a smaller dot difference would be required to maintain equal performance with typically developed participants. Here, however, we found no dot difference for equivalent performance between the groups.

The WCC hypothesis posits that a limited ability to process the global elements of a scene underlies the central cognitive disturbance in ASD. While there is no explicit local/global manipulation in the present study, the WCC implies reduced “top-down” appraisal of sensory inputs which arguably could involve an impairment of second-order judgements about first-order perceptual processes as we have demonstrated here. Additionally we note that reduced metacognitive efficiency for *non-social perceptual* judgements, may signal an unlikely link to social theories of ASD, e.g. Theory of Mind 29, since mind reading and metacognition both require meta-representation. In the case of mind reading we attribute mental states to others to better control our own social interactions, in the case of metacognition we monitor our own cognition, enabling us to learn more flexibly. Problems with meta-representation, then, may cut across social and non-social symptoms in ASD 4 and interventions targeting metacognitive ability may lead to improved outcomes in sensory and social functioning.

The present study has shown reduced metacognitive efficiency in ASD in the context of maintained perceptual performance. This was by design, as the first study to demonstrate isolated differences in metacognition in the absence of performance confounds 9. However, real-world perceptual decisions take place in dynamic and changeable environments where the objective uncertainty associated with the sensory evidence is liable to change. A key tenet of the Predictive Coding account of autism is difficulties with updating and estimating precision in the face of environmental change (volatility)5,7. Recent empirical studies have shown a tendency to over-learn about volatility in autism, which reduces confidence in prior expectations and also leads to increased phasic NA release (as indicated by pupil size; 11). Follow up work directly blocking the action of NA, showed that volatility-linked learning is attenuated30. Learning when we can learn and knowing when we know, are both second-order cognitive processes, so previously reported meta-learning difficulties in ASD align well with the findings of diminished metacognitive inefficiency reported here. In further support of this association, previous studies have shown that subjective confidence judgements track the inferential uncertainty in volatile probabilistic environments 13. Reference points (criteria) for confidence judgements need to be learned over time as evidence is accrued about task difficulty and there is suggestion that the role of the frontal cortex in metacognition lies in these higher order aspects of learning 8. One proposal is that the PFC, in concert with the anterior cingulate cortex (ACC), forms a meta-learning network, where the ACC signals volatility 31, and the PFC ratifies corresponding changes in confidence 32, enacted throughout task-relevant sensory cortex via neuromodulators that change cortical gain. We speculate that differences in this system may be apparent in ASD, but empirical studies are necessary to fully test this idea.

In summary, our current findings show for the first time that in ASD there are differences perceptual metacognition and its neuromodulatory correlates, which will have implications for meta-learning as hypothesised by Predictive Coding accounts of ASD 4,7. Future neuroimaging studies of confidence-linked learning will be necessary to address the neuroanatomical aspects of this proposal at the network level.

**References**

1. Geschwind DH. Genetics of autism spectrum disorders. *Trends Cogn Sci*. 2011;15(9):409-416.

2. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (5th Ed.)*. Washington, DC: Author; 2013.

3. Pellicano E, Burr D. When the world becomes ‘too real’: a Bayesian explanation of autistic perception. *Trends Cogn Sci*. 2012;16(10):504-510.

4. Friston KJ, Lawson R, Frith CD. On hyperpriors and hypopriors: comment on Pellicano and Burr. *Trends Cogn Sci*. 2013;17(1). doi:10.1016/j.tics.2012.11.003

5. Lawson RP, Rees G, Friston KJ. An aberrant precision account of autism. *Front Hum Neurosci*. 2014;8. doi:doi:10.3389/fnhum.2014.00302

6. Van de Cruys S, Evers K, Van der Hallen R, et al. Precise minds in uncertain worlds: Predictive coding in autism. *Psychol Rev*. 2014;121(4):649.

7. Palmer CJ, Lawson RP, Hohwy J. Bayesian Approaches to Autism: Towards Volatility, Action, and Behavior. *Psychol Bull*. 2017;143(5):521-542.

8. Meyniel F, Sigman M, Mainen ZF. Confidence as Bayesian Probability: From Neural Origins to Behavior. *Neuron*. 2017;88(1):78-92. doi:10.1016/j.neuron.2015.09.039

9. Fleming SM, Lau HC. How to measure metacognition. *Front Hum Neurosci*. 2014;8(433). doi:10.3389/fnhum.2014.00443

10. Rouault M, Seow T, Gillan CM, Fleming SM. Psychiatric symptom dimensions are associated with dissociable shifts in metacognition but not task performance. *Biol Psychiatry*. 2018;84(6):443-451.

11. Lawson RP, Mathys C, Rees G. Adults with autism overestimate the volatility of the sensory environment. *Nat Neurosci*. 2017;20(9):1293.

12. Hainguerlot M, Vergnaud J-C, de Gardelle V. Metacognitive ability predicts learning cue-stimulus associations in the absence of external feedback. *Sci Rep*. 2018;8(1):5602. doi:10.1038/s41598-018-23936-9

13. Meyniel F, Dehaene S. Brain networks for confidence weighting and hierarchical inference during probabilistic learning. *Proc Natl Acad Sci*. 2017;114(19):E3859. doi:10.1073/pnas.1615773114

14. Parr T, Rees G, Friston KJ. Computational neuropsychology and Bayesian inference. *Front Hum Neurosci*. 2018;12:61.

15. Martins ARO, Froemke RC. Coordinated forms of noradrenergic plasticity in the locus coeruleus and primary auditory cortex. *Nat Neurosci*. 2015;18(10):1483-1492.

16. Lempert KM, Chen YL, Fleming SM. Relating Pupil Dilation and Metacognitive Confidence during Auditory Decision-Making. *PLOS ONE*. 2015;10(5):e0126588. doi:10.1371/journal.pone.0126588

17. Allen M, Frank D, Schwarzkopf DS, et al. Unexpected arousal modulates the influence of sensory noise on confidence. *Elife*. 2016;5:e18103.

18. Joshi S, Li Y, Kalwani RM, Gold JI. Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron*. 2016;89(1):221-234.

19. Marshall L, Mathys C, Ruge D, et al. Pharmacological Fingerprints of Contextual Uncertainty. *PLOS Biol*. 2016;14(11):e1002575. doi:10.1371/journal.pbio.1002575

20. Hauser TU, Allen M, Purg N, Moutoussis M, Rees G, Dolan RJ. Noradrenaline blockade specifically enhances metacognitive performance. *eLife*. 2017;6.

21. Rouault M, Seow T, Gillan CM, Fleming SM. Psychiatric Symptom Dimensions Are Associated With Dissociable Shifts in Metacognition but Not Task Performance. *Biol Psychiatry*. Published online in press. doi:10.1016/j.biopsych.2017.12.017

22. Fleming SM, Ryu J, Golfinos JG, Blackmon KE. Domain-specific impairment in metacognitive accuracy following anterior prefrontal lesions. *Brain*. 2014;137(10):2811-2822.

23. Maniscalco B, Lau H. Signal Detection Theory Analysis of Type 1 and Type 2 Data: Meta-d′, Response-Specific Meta-d′, and the Unequal Variance SDT Model. In: *The Cognitive Neuroscience of Metacognition*. Springer; 2014:25-66.

24. Fleming SM. HMeta-d: hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. *Neurosci Conscious*. 2017;2017(1):nix007.

25. Kanai R, Komura Y, Shipp S, Friston K. Cerebral hierarchies: predictive processing, precision and the pulvinar. *Philos Trans R Soc B Biol Sci*. 2015;370(1668):20140169. doi:10.1098/rstb.2014.0169

26. Moran RJ, Campo P, Symmonds M, Stephan KE, Dolan RJ, Friston KJ. Free Energy, Precision and Learning: The Role of Cholinergic Neuromodulation. *J Neurosci*. 2013;33(19):8227-8236.

27. Happé F, Frith U. The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *J Autism Dev Disord*. 2006;36(1):5-25.

28. Mottron L, Dawson M, Soulieres I, Hubert B, Burack J. Enhanced perceptual functioning in autism: an update, and eight principles of autistic perception. *J Autism Dev Disord*. 2006;36(1):27-43.

29. Baron-Cohen S, Leslie AM, Frith U. Does the autistic child have a “theory of mind”? *Cognition*. 1985;21(1):37-46.

30. Lawson RP, Bisby J, Nord CL, Burgess N, Rees G. The Computational, Pharmacological, and Physiological Determinants of Sensory Learning under Uncertainty. *Curr Biol*. 2021;31(1):163-172.e4. doi:10.1016/j.cub.2020.10.043

31. Behrens TE, Woolrich MW, Walton ME, Rushworth MF. Learning the value of information in an uncertain world. *Nat Neurosci*. 2007;10(9):1214-1221.

32. Fleming SM, Dolan RJ. The neural basis of metacognitive ability. *Phil Trans R Soc B*. 2012;367(1594):1338-1349.

**Author Contributions**: RPL, KF and GR conceived the study. AJ collected the data. RPL analysed the data. RPL, KF and GR wrote the paper.

**Acknowledgements:** We would like to thank all the participants who gave up their time to take part in this research, Steve Fleming for proposing that we use this task and providing comments on an earlier version of this manuscript, and Chris Frith for the initial discussions that led to our hypotheses. This research was funded by a Wellcome Trust Senior Clinical Research Fellowship (100227; GR). RPL is supported by a Royal Society Wellcome Trust Henry Dale Fellowship (206691).

**Figure Legends:**

Figure 1: **Task and measures**. A) A typical trial, participants made a two-alternative forced choice judgement about which of two briefly presented circles contained more dots, followed by confidence ratings about that decision. For details see the main text and Supplementary Materials b) trial-by-trial measures of accuracy and confidence from a representative participant.

Figure 2: **Multiple indices of maintained perceptual performance**. There were no differences between the ASD and NT groups on measures of A) overall accuracy, B) task difficulty required to maintain equivalent accuracy, C) overall confidence ratings or, D) perceptual (type 1) sensitivity. Data points depict individual participants, blue and yellow shaded regions depict the standard deviation, light shaded regions depict the 1.96 standard error of the mean (95% confidence intervals) and the red line depicts the respective group mean for each measure

Figure 3: **Measures of metacognition**. A) Depicts reduced metacognitive sensitivity (meta-d’) in the ASD group and B) shows that metacognitive efficiency (meta-d’/d’) is significantly attenuated in the ASD group and significantly below optimal (dotted line). Data points depict individual participants, blue and yellow shaded regions depict the standard deviation, light shaded regions depict the 1.96 standard error of the mean (95% confidence intervals) and the red line depicts the respective group mean for each measure. D, E) Group estimated posterior for metacognitive efficiency (meta-d’/d’) shows significantly diminished metacognitive efficiency in the ASD group, consistent with the individual participant fits above.

Figure 4: **Confidence-linked metrics of noradrenergic function**. A) The effect of confidence on pupil size around the time of the dots stimulus (vertical dotted line indicates when the stimulus appeared). B) The effect of confidence on pupil size around the time of the confidence judgement (vertical dotted line indicates when the rating scale appeared). Blue solid horizontal lines indicate significant time clusters when the confidence-linked pupil response differed from zero in ASD participants (i.e., a significant regression of pupil size on trial-specific confidence ratings). The yellow solid horizontal line indicates the same for NT participants, and the black solid horizontal line shows time clusters where the associated regression slopes were significantly different in ASD compared with NT participants (2,000 permutations; cluster *α* = 0.05). In all pupil analyses this regression approach controlled for other explanatory and nuisance variables on a trial-by-trial basis (RT to make the confidence judgement, RT to make the initial perceptual decision, difficulty (∆ dots), and stimulus side (left, right)).

Figure 1

Figure 2



Figure 3



Figure 4



1. Although we abide by the terminology of the diagnostic and statistical manual (DSM-5) we wish to acknowledge that the term ‘autistic person’ is preferred by many people on the spectrum. [↑](#endnote-ref-1)