i-Comment

Weak priors versus overfitting of predictions in autism: Reply to Pellicano and Burr (*TICS*, 2012)

Sander Van de Cruys

Department of Experimental Psychology, University of Leuven, Leuven, Tiensestraat 102, Belgium; E-mail: sander.vandecruys@ppw.kuleuven.be

Lee de-Wit

Department of Experimental Psychology, University of Leuven, Leuven, Tiensestraat 102, Belgium; E-mail: lee.dewit@ppw.kuleuven.be

Kris Evers

Department of Experimental Psychology, University of Leuven, Leuven, Tiensestraat 102, Belgium; E-mail: kris.evers@ppw.kuleuven.be

Bart Boets

Department of Experimental Psychology, University of Leuven, Leuven, Tiensestraat 102, Belgium; E-mail: bart.boets@ppw.kuleuven.be

Johan Wagemans

Department of Experimental Psychology, University of Leuven, Leuven, Tiensestraat 102, Belgium; E-mail: johan.wagemans@ppw.kuleuven.be

Abstract. Pellicano and Burr (2012) argue that a Bayesian framework can help us understand the perceptual peculiarities in autism. We agree, but we think that their assumption of uniformly flat or equivocal priors in autism is not empirically supported. Moreover, we argue that any full account has to take into consideration not only the nature of priors in autism, but also how these priors are constructed or learned. We argue that predictive coding provides a more constrained framework that very naturally explains how priors are constructed in autism leading to strong, but overfitted, and non-generalizable predictions.

The regularities in the input tracked by Bayesian perception are constantly changing. A full explanation of autism has to explain not only how priors might differ in autism, but also how priors can be dynamically adjusted to a changing environment. This makes predictive coding a particularly attractive implementation of Bayesian perception. Predictive coding replaces the priors in Bayesian reasoning by top-down predictions about what sensory input to expect next in the current context, while Bayesian sensory evidence is replaced by prediction errors, i.e. that part of the input that has not been predicted by higher levels. These prediction errors are crucial for learning because they signify that a particular top-down prediction is not optimally tuned to the regularities in a particular environment. Prediction errors therefore trigger learning, such that today's prediction errors shape tomorrow's (hopefully more optimal) predictions. Note that no two experiences are completely identical; thus, any new experience will inevitably contain some prediction error. It is possible, however, for the system to tolerate some degree of prediction error through a context-sensitive adjustment of precision. The precision of prediction errors can determine the relative weight of top-down versus bottom-up information (Friston, 2010). Optimally, precision should be increased in situations where there still are learnable regularities (not random noise) and reduced when there are not. This context-dependent adjusting of precision constitutes the mechanism underlying attention in a predictive coding framework (Friston, 2010; Kok, Rahnev, Jehee, Lau, & de Lange, 2012). Thus, precision—and the extent to which prediction errors are upregulated to update future predictions—is not given; it has to be estimated based on reliability and relevance of predictive cues. We propose that in autism, the flexible adjusting of precision is lacking.

A chronically increased precision will initiate new learning at every new instance. Hence, future predictions are shaped by noise or contingencies that are unlikely to repeat in the future (also called overfitting). An organism that developed this kind of priors will have strong predictions on what to expect next but such predictions will quasi-never be applicable. As a result, every new instance will be treated as an "exception," incomparable to previous experiences (little or no generalization). Because of the strong, specific predictions, the organism will be flooded by strong, specific discrepancies (prediction errors). Therefore, it may seem that top-down predictions are weaker ("hypopriors") in autism because prior knowledge seems to influence perception less (Mitchell & Ropar, 2004), but according to our view this stems from the fact that predictions are too narrowly tuned to the (noise in the) world (Feldman, 2013).

Our proposal complements a commentary by Friston, Lawson, and Frith (2013), who argued that predictive coding offers a neurally plausible implementation of Bayesian inference. In addition to its potential neural implementation in the hierarchical human visual system, we want to emphasize that a deficit in the flexible adjustment of precision can also explain how priors are generated as overfitted predictions. Our account is also complementary to a commentary on Pellicano and Burr (2012) by Brock (2012), who highlighted that the behavioral evidence associated with weak priors could in fact result from an overweighting of sensory input, rather than a reduction in top-down prediction. Our interpretation could be described as the "worst of both worlds," there is both an overprecision at early sensory stages of processing, and this leads to the upregulation of prediction errors that induce priors that are too narrow. The account above has several neural and behavioral consequences which are consistent with what is found in autism. *First*, mismatch negativity studies measuring ERP responses to deviant (unexpected) stimuli confirm that people with autism are perfectly able to form priors or predictions, and that their reaction to (small) violations of these predictions is even faster or stronger (Ferri et al., 2003; Kujala et al., 2007). *Second*, because of strong, specific prediction errors, processing will be stuck in lower levels of the perceptual hierarchy. Again, predictions are generated but they remain more specific lower level predictions, rather than global, higher level predictions. This is reflected by numerous studies showing improved local, detail-oriented performance and impaired global, configural processing in autism (Dakin & Frith, 2005). Note that this also offers a potential explanation for inconsistencies in the literature on the lack of global processing in autism. According to this view, people with autism can generate more global, higher level predictions, but they will often not be able to apply them. *Third*, problems in executive and perceptual flexibility in patients with autism (Ciesielski & Harris, 1997) also suggest they do have strong, specific priors (expectations) about task rules and perceptual input. If one assumes that the establishment of a specific task set, or the representation of rules for guiding behavior in a specific task, can also be understood in terms of the representation of priors, then one should expect that weak or broad priors should make switching between sets in an executive task easier, not harder. With regard to perceptual switching in bistable figures, little would be needed to shift perception to another stable interpretation if only weak priors were developed. This does not seem to be the case in autism (Ciesielski & Harris, 1997), although more research on bistable perception is needed. *Fourth*, the increased prevalence of "savant" abilities in autism (Mottron, Dawson, Soulieres, Hubert, & Burack, 2006) is more easily explained within an account that assumes strong, specific priors, that could enable one to make very specific expectations, rather than with an account that starts from weaker priors. These special talents seem to be grounded in a better discrimination of perceptual differences (e.g., pitch).

In daily life, autistics insistence on sameness and their acute awareness of minute changes (prediction errors) also seem to speak to the view proposed here. The taxing experience in autism (cf. sensory overload) may result from a perceptual system that continuously signals prediction errors, indicating that there always remains something to be learnt still and that attentional resources are needed. The accompanying negative feelings could cause these patients to avoid the most variable or unpredictable situations where context-dependent high-level predictions are more important than concrete perceptual details. This may be the case for social interaction in particular. The overwhelming prediction errors cause these patients (or their caregivers) to externalize and enforce predictability through exact routines and patterns in their daily activities.

In summary, we are grateful that Pellicano and Burr (2012) ignited the discussion because we agree with them that it is useful to try to understand the perceptual changes in autism in terms of changes to Bayesian priors. However, we pointed to several sources of evidence to suggest that rather than having uniformly weak priors, people with autism often develop very strong priors, or expecta-

tions, in particular contexts. We think 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 particular, we think autism is associated with an inability to flexibly adjust the degree of precision in a different context. The increased precision will cause more prediction error to be upregulated and induce learning. This will lead to overfitting to specific contexts, and priors that may sometimes be very strong, but which are unlikely to be generally applicable.

References

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

Ciesielski, K. T., & Harris, R. J. (1997). Factors related to performance failure on executive tasks in autism. *Child Neuropsychology, 3*(1), 1–12. doi:10.1080/09297049708401364

Dakin, S., & Frith, U. (2005). Vagaries of visual perception in autism. *Neuron*, *48*(3), 497–507. doi:10.1016/j.neuron.2005.10.018

Feldman, J. (2013). Tuning your priors to the world. *Topics in Cognitive Science, 5*(1), 13–34. doi:10.1111/tops.12003

Ferri, R., Elia, M., Agarwal, N., Lanuzza, B., Musumeci, S. A., & Pennisi, G. (2003). The mismatch negativity and the P3a components of the auditory event-related potentials in autistic low-functioning subjects. *Clinical Neurophysiology*, *114*(9), 1671–1680. doi.org/10.1016/S1388-2457(03)00153-6

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

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

Kok, P., Rahnev, D., Jehee, J. F., Lau, H. C., & de Lange, F. P. (2012). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex*, *22*(9), 2197–2206. doi:10.1093/cercor/bhr310

Kujala, T., Aho, E., Lepistö, T., Jansson-Verkasalo, E., Nieminen-von Wendt, T., von Wendt, L., & Näätänen, R. (2007). Atypical pattern of discriminating sound features in adults with Asperger syndrome as reflected by the mismatch negativity. *Biological Psychology*, *75*(1), 109–114. doi:10.1016/j.biopsycho.2006.12.007

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

Mottron, L., Dawson, M., Soulieres, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders*, *36*(1), 27–43.

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