

# Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity

Bolis, D.<sup>a, b, c\*</sup>, Balsters, J.<sup>c</sup>, Wenderoth, N.<sup>c</sup>, Becchio, C.<sup>d, e</sup>, Schilbach, L.<sup>a, b, f, g</sup>

<sup>a</sup> Independent Max Planck Research Group for Social Neuroscience, Max Planck Institute of Psychiatry, Kraepelinstrasse 2-10, 80804 Munich-Schwabing, Germany

<sup>b</sup> International Max Planck Research School for Translational Psychiatry (IMPRS-TP), Munich, Germany

<sup>c</sup> Neural Control of Movement Lab, Department of Health Sciences and Technology, ETH Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

<sup>d</sup> Department of Psychology, University of Torino, Torino, Italy

<sup>e</sup> Robotics, Brain and Cognitive Sciences, Fondazione Istituto Italiano di Tecnologia, Genova, Italy

<sup>f</sup> Department of Psychiatry, Ludwig Maximilian Universität, Munich, Germany

<sup>g</sup> Outpatient and Day Clinic for Disorders of Social Interaction, Max Planck Institute of Psychiatry, Kraepelinstrasse 2-10, 80804 Munich-Schwabing, Germany

## Abstract

Drawing on socio-cultural theories and Bayesian accounts of brain function, in this article we construe psychiatric conditions as disorders of social interaction, to fully account for their complexity and dynamicity across levels of description and temporal scales. After an introduction of the theoretical underpinnings of our integrative approach, we take Autism Spectrum Conditions (ASC) as a paradigm example and discuss how neurocognitive hypotheses can be translated into a Bayesian formulation, i.e. in terms of predictive processing and active inference. We then argue that consideration of individuals (even within a Bayesian framework) will not be enough for a comprehensive understanding of psychiatric conditions and consequently put forward the *dialectical misattunement hypothesis*, which views psychopathology, not merely as disordered function within single brains, but as a dynamic interpersonal mismatch that encompasses various levels of description. Moving from a mere comparison of groups, i.e. 'healthy' persons versus 'patients', to a fine-grained analysis of social interactions within dyads and groups of individuals will open new avenues and may allow to avoid an overly neurocentric scope in psychiatric research, as well as help to reduce social exclusion.

**Keywords:** autism; dialectical misattunement; social interaction; intersubjectivity; cultural historical activity theory; enactivism; predictive processing / coding; active inference; dialectical approach; Vygotsky; Bayes.

\* Corresponding author at: Independent Max Planck Research Group for Social Neuroscience Max Planck Institute of Psychiatry Kraepelinstr. 2-10, 80804 Munich, Germany. E-mail address: [dimitris\\_bolis@psych.mpg.de](mailto:dimitris_bolis@psych.mpg.de)

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

**Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press**

*δὶς ἐς τὸν αὐτὸν ποταμὸν οὐκ ἂν ἐμβαίης.*

*You could not step twice into the same river.*

*τὰ ὄντα ἰέναι τε πάντα καὶ μένειν οὐδέν.*

*All flows, nothing stays.*

**Heraclitus (ca. 535 - ca. 475 BC)**

*Through others, we become ourselves.*

**Lev Vygotsky (1896-1936)**

## 1. A synthesis of dialectical and computational perspectives

### 1.1 Psychiatry through a dialectical lens

In this paper, we will put forward an integrative approach for revisiting psychiatric conditions, taking *dialectics* as a point of departure. The latter could be considered as an evolving school of thought, met in various historical contexts (e.g. Greek dialectic, Chinese dialectic, Hegelian dialectic, Marxian dialectic [1,2]) critical to both reductionism and dualism. It asserts that phenomena cannot be meaningfully understood by reducing them into single levels of description (cf. reductionism) or assuming a metaphysical independence between levels (cf. dualism), but should be rather studied in their wholeness, inner contradiction and movement. In this light, human mind and psychopathology cannot be understood in isolation from society, the body and social interaction. To quote Hegel “*to know, or, in other words, to comprehend an object is equivalent to being conscious of it as a concrete unity of opposed determinations*” [3,4]. We will, therefore, try to overcome traditional dichotomies, such as organism/environment, by viewing them as both a result and a cause of reciprocal adjustments, or individual/society by considering the whole and the part as, albeit partially autonomous, highly interdependent levels of organization. In this effort, we will also draw upon accounts of intersubjectivity, which emphasize that single levels of analysis or cutting off the part from the whole may severely limit our understanding of a phenomenon. We will emphasize viewing psychiatric conditions, not as static conditions driven by a single cause, but rather as the outcome of an interplay of multiple and diverse factors (**Fig. 1**) and to be more specific as a process of circular causality among different levels of description (e.g. biological, cognitive-behavioral and socio-cultural), as well as multiple functions within a level (e.g. action and perception within the cognitive-behavioral level), unfolding over different temporal frames (e.g. evolutionary, cultural, social, individual-psychological, subindividual-biological developing scales; based on Lev Vygotsky’s and colleagues’ views on human development [5,6]).<sup>1</sup>

---

<sup>1</sup> Please note the specific definition and distinction between levels, functions and temporal frames, as put forward here, are made for intelligibility purposes only and it should not be taken as implying dichotomies; processes and their interrelationships appear complex, continuous and overlapping in reality (e.g. see [145]).

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

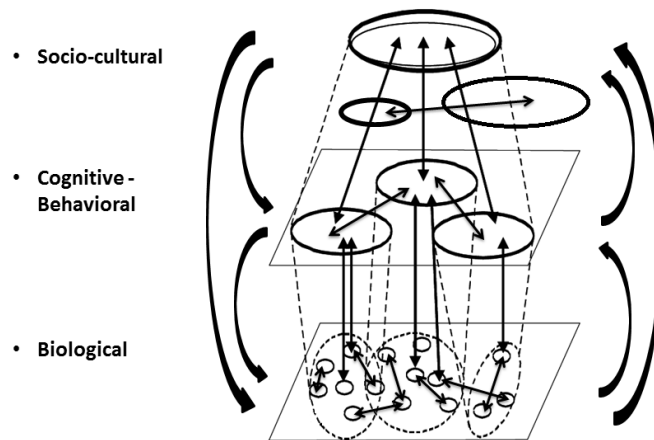
**Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press**

Indeed, contrary to a common assumption that a full description on a micro-spatiotemporal level is causally complete, it has been suggested that a genuine causal emergence on a macro level might also be possible [7]. Importantly, such an emergence is not to be solely attributed to a weakness of experimental means to fully grasp the micro-phenomena, but rather due to inherent characteristics of systemic processes themselves. For example, coarser mechanisms on a higher level might appear more robust in terms of causality, than relevant stochastic micro-processes. Thus, a genuine causal emergence on a macro level is necessary for a complete description. In fact, this is a conclusion from physics where the circular causality between the microscopic and the macroscopic is well established in terms of concepts such as the slaving principle and the center manifold theorem. In brief, these theorems suggest the emergent macroscopic (order parameters) that describe the whole enslave the microscopic components that constitute the whole. This induces a circular causality that lies at the heart of synergetics (cf. [8]). It also speaks to the circular causality to which enactivism and embodied (situated) cognition approaches appeals. Following such a line of thought this paper will argue that while considering neurobiological and phenomenological processes is an important step toward the understanding of psychiatric conditions, it may remain incomplete as further levels of analysis, such as sociocultural processes and generally social structure, are neglected. For instance, structures promoting social exclusion or competitiveness, as opposed to communication and collaboration could distinctly shape individual behavior, mental reality and biological mechanisms. Here, our approach heavily leans on work from the cultural historical activity theory, which re-interpreted human development across a variety of conditions, as a dynamic interplay between biological and sociocultural forces ([6,9,10] on the work of Lev Vygotsky and colleagues). Notably, the aforementioned variety of conditions were not limited to what one could think of ‘social conditions’, but rather included individuals who were both deaf and blind, to give an example. The organic condition can of course still affect the construction of the social self via atypical development, if amelioration of social exclusion is not taken into account. As Vygotsky, pointed out:

*“the confusion and failure to differentiate the organic from the cultural, the natural from the historical and the biological from the social [...], inevitably leads to a fundamentally incorrect understanding and interpretation of the data (observations)”* (excerpt from Vygotsky’s work; translated in [6]).

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

***Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press***



**Figure 1.** Schematic depiction of dynamic interrelationships: between multiple levels (e.g. biological, cognitive-behavioral, sociocultural) and functions (e.g. including but not limited in the functionality of multiple neuromodulators or bacteria on the 1<sup>st</sup> level, body-/neuro-systemic and phenomenological aspects on the 2<sup>nd</sup> level, social structure, institutions or cultural practices on the 3<sup>rd</sup> level), interacting in several temporal scales. Please note arrows may appear static on the image, but we interpret them as representations of developing interrelationships reflecting both quantitative and qualitative changes (cf. dialectics). Certain additional core levels of description, i.e. the (micro-/macro-) physical levels, have been omitted from this illustration.

## 1.2 Psychiatry through a computational lens

In our effort to adopt an integrative perspective, we will use Bayesian accounts of cognition and behavior as powerful tools of analysis within the level of the individual, but most importantly we will suggest ways of going beyond the individual as the unit of analysis, and eventually overcoming limitations of a single-level approach (see Sections 5 and 6). Computational psychiatry can be thought of as lying on the interface between computational neuroscience and clinical psychiatry. It deploys computational (e.g. Bayesian) modeling in order to mechanistically describe psychiatric conditions (e.g. [11,12]). A more specific hierarchical Bayesian approach to perception and action, which we will focus on here, has been described as the *predictive coding* (also mentioned as predictive processing; a term which we will be using in this article) and *active inference* account. In brief, according to such a perspective, the brain's ultimate goal is the long-term minimization of free energy, which (as we will explain later, under simplifying assumptions) can be thought of as the "prediction error", i.e. the discrepancy between incoming information and generated predictions, based on consolidated experience. Importantly, this is thought to be accomplished through two main avenues, namely either via updating the beliefs one holds for aligning them with the environment (i.e. predictive processing), or through action, which can help to experience the environment in accordance with prior beliefs (i.e. active inference). Here, it should be noted, that Bayesian beliefs inherent in any Bayesian approach to cognition should largely thought of as subpersonal. In other words, the experience subtended by predictive processing is not necessarily a conscious experience but more like a percept (or possibly a causative experience; i.e. qualia), embracing also other 'automatic' processes such as homeostatic

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

**Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press**

control. One of the many interesting aspects of this account is that perception, learning and action are not considered as isolated and passive processes, but they constitute interconnected processes, which an organism actively deploys for making sense (or to put it in computational terms, ‘model’) the world, in order to maintain its current living form [13].

### 1.3 The dialectical misattunement hypothesis of social interaction and a Bayesian account of intersubjectivity

Taken together, we suggest that formally considering, both quantitative and qualitative, dynamically changing interrelationships between and within levels of description (**Fig.1**), as well as temporal scales will be essential for a comprehensive understanding of complex psychiatric conditions, such as ASC. In light of this, the purpose of this paper will be three-fold: Firstly, to consider the integration of diverse within-level (i.e. neurocognitive) processes embedded in a common framework, i.e. the predictive processing and active inference account. Secondly, to outline the importance of taking into account interrelationships across levels (i.e. the individual and the collective), via putting forward the misattunement hypothesis of social interaction. Thirdly, to ultimately motivate the development of a Bayesian account of intersubjectivity rather than of individual brains. Importantly, we also highlight the practical implications of our theoretical approach (i.e. ethical, research, clinical and pedagogical). Taking Autism Spectrum Conditions (ASC) as a paradigm case, we will give a description of the general framework of our approach. More concretely, we will first review the field of autism research, with an emphasis on recent interest in providing a Bayesian formulation of ASC. Based on this, we will argue in favor of adopting the Bayesian accounts of brain function as a framework to integrate seemingly contradictory neurocognitive hypotheses. Then, we will discuss different accounts of intersubjectivity, which share a common ground by stating that individual level-analyses do not suffice for a comprehensive understanding of social perception and cognition. Bringing together a dialectical perspective to human communication and Bayesian (i.e. predictive processing and active inference) accounts of individual mechanisms [14], we will introduce the *dialectical misattunement hypothesis* of social interaction, which emphasizes the interdependence of individual and collective levels of description.

More concretely, the dialectical *misattunement hypothesis* rethinks ASC, not merely as a disorder of the individual brain, but as cumulative misattunement between persons as well. Misattunement across persons can be thought of as disturbances of the dynamic and reciprocal unfolding of an interaction across multiple time scales, resulting in increasingly divergent prediction and (inter-)action styles. Consequently, with regards to neuroscientific research we propose moving from focusing only on comparing groups of *individuals* to considering types of interaction *between persons* (e.g. homogeneous dyads consisted of either only neuro-typical persons or only persons with a certain condition, as well as heterogeneous dyads; including both tuned and non-tuned interactions<sup>2</sup>). Here the hypothesis holds clear predictions: Interactions within homogeneous dyads are expected to appear smoother, compared

---

<sup>2</sup> The term ‘tuned’ here refers to multiple aspects: tuning expectations of either or both the interactors, as well as facilitating the interaction via tuning the communication medium (e.g. social conventions, as well as the cultural or technological environment, which the interaction is embedded in).

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

**Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press**

to heterogeneous dyads. Additionally, tuned interactions of either homogeneous or heterogeneous dyads should appear as most effective. If these hypotheses are valid, the definition of a psychiatric condition as ASC, can be thought of as relative to the ‘other’, and generally the social context. Such an approach, will eventually allow us to escape an overly neurocentric research scope in psychiatry. Along similar lines, we suggest that clinical and pedagogical practices should move beyond the individual, to monitoring, evaluating and facilitating processes at the interpersonal level. Also, re-viewing ASC as a misattunement between people, and not as disorder of the brain *per se*, may help to alleviate social stigma and reduce social exclusion.

We will end by outlining a Bayesian account of intersubjectivity, referred to as the “*observing-the-interactors*” scheme, which will allow us to computationally describe the interplay of individual and collective levels of activity during social interactions. Subsequent papers will delineate a practical approach for testing the misattunement hypothesis of social interaction, based upon hierarchical models of interpersonal interactions [15] and two person psychophysiology [16]. In what follows, we focus on autism, but the proposed approach more generally applies to any process evolving at the interface between the intra- and the inter-personal, including social exclusion across different conditions.

## 2. Traditional views on Autism Spectrum Conditions

Although sparse references about resembling cases may have existed before [17], it was not until the 1940s that Hans Asperger and Leo Kanner described the condition of autism. Today autism is considered as a neurodevelopmental disorder spanning a spectrum, which is characterized by impairments in social interaction and communication, as well as restricted, repeated behaviors and interests. It is also not uncommon for individuals with an autism spectrum condition to show enhanced abilities for specific cognitive aspects including perception [18], attention [19] and memory [20]. While some approaches have focused on the impairments, other accounts encompass both impaired and enhanced skills [21,22], especially when it comes to the so-called “high-functioning” end of the spectrum. In the past half century, a number of different cognitive hypotheses have been pursued in order to understand core aspects of ASC. Although several important ideas have helped to shed light on specific facets, there is still no consensus about a single theory that could offer a universal and yet specific explanation of the condition. Here we will primarily focus on the “five big ideas” about autism, as suggested by Uta Frith [23]:

Firstly, Baron-Cohen and colleagues proposed that individuals with ASC lack a specific meta-representational capacity, namely a “*theory of mind*” (*ToM*), which prevents them from inferring upon other people’s mental states [24]. As a consequence of this, individuals with ASC - so it is assumed - cannot know about other people’s beliefs, emotions, desires, perceptions and intentions. In light of findings, that individuals with ASC can make a conscious effort to think about others’ mental states, it has been suggested that implicit, namely spontaneous mechanisms of mentalizing might be the ones that are primarily linked to relevant difficulties in ASC, rather than explicit ones, which might be easier compensated for through learning [25,26].

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

**Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press**

The second big idea focuses on a special category of neurons, the so called “mirror neurons” [27,28], which are active both when an action is performed and observed. The *broken mirror neuron (BMN)* hypothesis proposes the explanation of impaired social skills in ASC on the basis of a dysfunctioning mirror neuron system (MNS) [29,30]. A number of studies offered supportive evidence to the involvement of MNS [29,31,32]. However, both the validity of a broken MNS and a direct, causal relationship between the MNS and social skills in ASC have been challenged by other reports [33,34]. Differences in MNS activation between neurotypical individuals and people with an ASC could be alternatively traced back to earlier modulatory effects of the mentalizing system as well (e.g. [35,36]).

Alternatively, the *social motivation (SM)* hypothesis focuses on motivational rather than ‘purely cognitive’ aspects [37]. It proposes that people with ASC lack the inherent social drive, which would assist them in exploiting the necessary learning opportunities for developing expertise in social cognition. More precisely, the hypothesis is settled upon the fact that social orienting, social seeking and liking, as well as social maintaining appear to be affected in ASC. On a biological level, the focus is placed on the human reward system, where either specific social impairments or more general reward-related dysfunction could explain the behavioral findings. A suboptimal oxytocin regulation has also been implicated in ASC, which could for example reflect differences in relating social stimuli to rewarding values (e.g. see [38–41]).

The fourth idea, namely the *weak central coherence (WCC)* hypothesis considers ASC as a different, detailed-oriented cognitive style [18,42–44]. More precisely, it claims that people with an ASC tend to process information locally, rather than globally. It predicts that people with ASC will have difficulties in perceiving information in context. According to this idea, people with ASC perceive the world differently in a number of aspects, including visual, auditory and linguistic functions. Later, the *enhanced perceptual functioning* hypothesis attributed this local bias to a superiority of detail processing *per se* and not due to inferiority of global information processing [45]. In the meanwhile, the *monotropism* hypothesis proposed a generalization from the tendency to focus on a local level to a need of focusing on a single source level of information [46].

Finally, the *executive dysfunction (ED)* hypothesis focuses on the difficulties, that people with ASC face when it comes to executive functions, i.e. problems primarily associated with functions such as planning, flexibility, inhibition and working memory [47–50]. For instance, difficulties related to dealing with novel situations and improvising, as well as perseverative stereotyped behavior in ASC can be explained by ED. This hypothesis has been taken to suggest that the study of frontal cortex function should be particularly relevant for a neurofunctional understanding of ASC.

To conclude this brief introduction of various accounts of ASC, it can be said that a number of different hypotheses have provided important insights into specific aspects of the condition of ASC; still, none of them is considered to provide a global explanation. In fact, it has been argued that a single explanation at the cognitive, neural or genetic level might be intractable [51–53]. However, an interest in a potentially unifying account has recently re-emerged while making reference to and drawing upon the

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

**Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press**

*Bayesian brain* hypothesis and particularly the *predictive processing and active inference* scheme [16,54–62]. It is to the discussion of this approach and its relevance for ASC that we now turn.

### 3. Bayesian approaches

#### 3.1 The Bayesian Brain Hypothesis

The main premise of the Bayesian brain hypothesis rests on the idea that the brain represents information accessed via the sensory organs in the form of probability densities, as opposed to single numbers, which are continuously updated, as if following a specific set of mathematical formulas based on the Bayes theorem. Crucially, this allows for optimal information integration both in time and space, multimodal cue integration, as well as flexible information manipulation without the need to commit to particular decisions at an early stage of processing [63]. To put it simply, through a Bayesian lens one can view the brain as an organ which calculates and maintains probabilities about events in the environment or about the self, via a combination of already gained experience and newly sensed information. Crucially, the more confidence (i.e. precision) is placed on the validity of experience (i.e. prior beliefs) the less the latter is updated in the face of new incoming information (i.e. evidence).

To make it more intuitive, let us imagine a young woman, Penelope, who living in Southern Greece, wakes up on a summer morning late for her work. The blinds are shut down and there is no time to check the weather outside the window. Will she take her umbrella on the way out? Based on her experience (i.e. prior beliefs; it rarely rains in Southern Greece in the summer), she decides not to take her umbrella with her. However, in the evening it happens to rain (evidence). The next day, Penelope, bringing together experience and the previous day's facts, thinks there might be a slightly higher probability of raining (i.e. posterior belief), but this is still not high enough to persuade her that carrying an umbrella might be a good idea. After several days of raining, she eventually decides to put the umbrella in her bag. She has come to believe that the probability of raining is high enough these days, despite her opposing experience of previous years. Perhaps not surprisingly from a Bayesian point of view, Penelope still keeps the umbrella with her for a few days after the weather has been sunny and dry again. Before concluding our example, it is worthwhile to introduce the concept of precision, which can be generally thought of as the confidence about a certain belief. Let us imagine a second scenario, where Penelope wakes up on a summer morning in Japan, where she has been travelling for a few days. She has heard that weather is generally dry in summer in the city she stays. Yet, on the first day it does happen to rain. Interestingly, already from the next day she decides to take an umbrella with her. Why did she change her mind so quickly in this case? Adopting a Bayesian perspective, one could argue that Penelope, although holding a high prior belief about not raining, changes her mind quickly due to the relatively low confidence (i.e. precision) she places on these prior beliefs of her, which have been the result of rumors and not her own experience.

#### 3.2 The hypo-priors hypothesis of autism

Coming back to our main example of ASC, Pellicano and Burr adopted a Bayesian standpoint to argue that non-social features of ASC might be well explained in reference to attenuated Bayesian priors (i.e.

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

***Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press***



priors of relatively low precision, so-called *hypo-priors*) [54]. This hypothesis anticipates a relatively more precise perception in ASC, driven primarily by perceptual evidence as opposed to prior knowledge, as well as the sense of being overwhelmed by this information, a common complaint of persons with ASC. Moreover, the hypo-priors hypothesis predicts the impedence of performance in ambiguous situations when prior knowledge is crucial for optimally solving a perceptual problem of inference. Finally, it was considered that a different learning style, namely one resembling *overfitting* in machine learning, and differences in adaptation can also be explained by this hypothesis (cf. [64]).

The hypo-priors hypothesis was then reformulated [56,57] within the predictive processing scheme, a more specific Bayesian account [65–68], while considering social aspects of *individual* cognition and behavior [60,61]. It is worth noting that the importance of difficulties related to predictions had been noted in the autism literature in the past as well (e.g. [69]). However, the more recent shift toward focusing on predictive processing and particularly on the concept of precision as described above can offer a potentially unifying explanation of autistic symptoms and directly relate computational findings with tractable neurobiological mechanisms. Before explaining how a predictive processing and active inference framework could, therefore, facilitate research into autism, we will first present the underlying basic ideas.

### 3.3 Predictive processing and active inference

The general idea of predictive processing and active inference is not new. For instance, one can find indications in Hermann von Helmholtz [70], who spoke about “unconscious inference” in the 19th century, drawing on ideas going back to ancient philosophers. Additionally, relevant traces can be found in ideas such as the reafference and ideomotor principles [71–73]. To put it simply, within a predictive processing and active inference framework the brain is essentially viewed as a “prediction machine” whose ultimate goal is the minimization of “prediction error” by deploying hierarchical generative models. More precisely, higher levels of a hierarchy continuously produce predictions, which are tested against the input information of the immediate lower levels. The discrepancy between predictions and incoming information, i.e. the “prediction error”, is propagated to higher-levels, reconfiguring the system to optimize its next predictions. Notably, propagating only the error and not the actual incoming information to higher levels is an efficient and resource-oriented way of reducing the bandwidth of the processed information, which is also exploited in data compression techniques, such as the common JPEG format. In short, two processes take place at the same time in opposite directions; predictions are propagated backward from higher to lower levels, trying to *explain away* prediction errors, and prediction errors are propagated forward from lower to higher levels, updating predictions (**Fig. 2**).<sup>3</sup> The hierarchical structure of the model is of immense importance because it enables the brain to optimize its own (empirical) priors on the fly. Additionally, it allows for effective representations of increasing abstraction. From a neurobiological perspective, forward connections may arise in superficial pyramidal

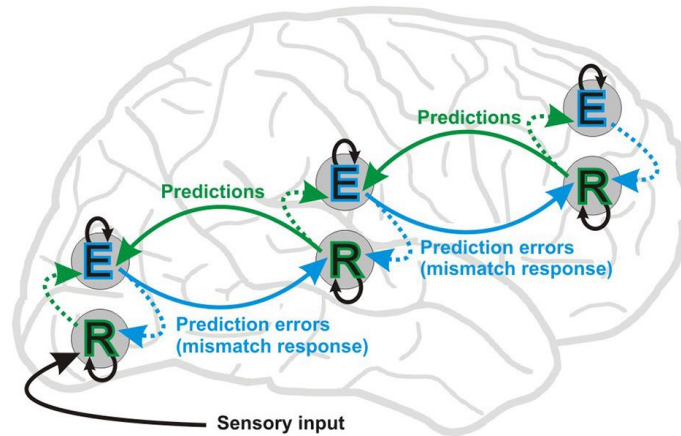
---

<sup>3</sup> Please note the new perspective, which is introduced with the predictive processing definitions of ‘backward’ and ‘forward’ connections, contrasted with the ‘feedback’ and ‘feedforward’ ones, since in the context of predictive processing the backward connections are the ones providing feedback via prediction error information on the forward stream of predictions [68].

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

**Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press**

cells, whereas the sources of backward connections are assumed to reside in deep pyramidal cells [74,75].



**Figure 2.** A simplified representation of the predictive processing idea (taken from: Stephanics *et. al*, 2013): representation units (R; deep pyramidal cells) receive inputs (blue arrows) from error units (E; superficial pyramidal cells) of the same (dotted line) and lower levels, while error units receive inputs (green arrows) from the same (dotted lines) and higher levels. Black arrows represent inhibitory intrinsic connections.

At this point, it is important to place the predictive processing in the more general context of active inference (a corollary of the free energy principle). Crucially, active inference takes predictive processing beyond the domain of perceptual inference and provides an account of action. The brain can be seen as inferring upon the causal structure of the world by updating “beliefs”, which are represented as probability densities. Most simply, the latter would take the form of Gaussian distributions, fully defined by their mean (i.e. expectation) and variance (i.e. inverse precision). Under this simplifying assumption (i.e. the *Laplace assumption*), the generalization of prediction error minimization to “free energy” minimization becomes mathematically more evident.<sup>4</sup> The latter then takes the form of a difference between the predictions of a model and the representations to be predicted [13]. Indeed, free energy had been originally formulated for confronting the difficult problem of exact inference, transforming it into an easy problem of optimization. It could be possible, that a similar trick is used by the brain in order to efficiently approximate the inference problem in a quasi-optimal Bayesian way. Interestingly, the *free energy* principle has been proposed as a potentially unifying brain theory, accounting for action, perception and learning. In short, an agent has two options for suppressing free energy; first by selectively sampling the environment for fulfilling its own expectations (i.e. through acting referred to as active inference) and second by optimizing these expectations for better matching with its sensations (i.e. through perception and learning referred to as predictive processing [76,77]). More broadly, one could sketch a path which, starting from the existence of life (as a process leading to a restricted number of states), passes through entropy (referring to a tendency to resist the 2nd law of thermodynamics), surprise (viewing entropy here as a mean value of surprise over time), free energy (as an upper bound of

<sup>4</sup> In this setting free energy can be regarded as an approximation, namely an upper bound, to Bayesian model evidence, which is the probability of observing the data given a specific model.

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

**Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press**

surprise) and eventually leads to prediction error, which as we pointed out can be considered as the free energy under certain simplifying assumptions. As provocatively put by Karl Friston “*the motivation for minimizing free energy has hitherto used the following sort of argument: systems that do not minimize free energy cannot exist*”.

Crucially, in the setting of predictive processing and active inference the degree of prediction updating (i.e. the learning rate) is controlled by the relative precision of successive levels. More precisely, it is proportional to a relative precision-weighted prediction error. This makes sense, since it would be generally desirable for an agent to update their beliefs first when the prediction error is large and second when they are unsure (low precision or confidence) about their prior beliefs compared to incoming information of lower levels in the hierarchy (about the importance of precision see [78]). Importantly, the idea of an updating rule proportional to the precision-weighted prediction error is a potentially neurobiologically plausible account, where precision is assumed to be represented by the gain of superficial pyramidal cells calculating precision errors [79–81]. Psychologically, increases and decreases in the precision of sensory prediction errors have been associated with sensory attention and attenuation respectively. In other words, attending to (or attenuating) a sensory stream is, under predictive processing, mediated by affording more (or less) precision to that stream [82].

Before concluding this introduction to predictive processing and active inference, it is worth noting that this scheme could be considered as a dialectical framework in and of itself. Firstly, it defines action and perception as the interplay between two closely intertwined avenues for minimizing prediction error. New perceptual states can inform future actions, while informed adjustment and sampling of the environment (i.e. action) decisively contributes to updating perception. Essentially perception and action become here two dialectical facets of the same process, namely minimization of free energy. Additionally, prediction updating and activity can be viewed as dialectical processes in time, between prior experience and incoming information, whose confrontation yields adjusted relations between environment and the self, either through updating current beliefs or the perceived environment itself. We again see here a circular causality that is central to enactive (Bayesian) inference - and speaks to related notions in enactivism and embodied cognition (see Section 5). After having provided a general introduction of the predictive processing and active inference framework, its putative role in understanding autism will be presented in the following.

### 3.4 The aberrant precision hypothesis of autism

It has been suggested that considering the role of precision in cognitive and behavioral processes could be important for understanding differences between neurotypical persons and people with an ASC: Indeed, there is preliminary neurobiological evidence, with regard to the functionality of certain neuromodulators that is suggestive of aberrant precision in ASC (see [60]). Additionally, several, psychological findings in ASC could be putatively attributed to aberrant precision estimation (e.g. see [61,83]). For instance, sound and visual stimuli hypersensitivity is typically observed in people with ASC (e.g. [45]). Through a predictive processing and active inference lens, consideration of irrelevant information due to increased precision can possibly lead to perceptual overload, or in other words perceptual hypersensitivity. Furthermore, stereotypies, repetitive behaviors and self-stimulation, all

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

**Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press**

commonly observed in ASC, could be viewed as efforts for creating scenarios of reduced prediction error, because other pathways fail to do so. Finally, another core attribute of ASC, i.e. withdrawal to one's own self, might constitute an alternative strategy of generally keeping prediction errors low. This kind of behavior could also be linked to an attenuation of motivational factors due to a persistent inefficiency to trigger reward through decreasing prediction errors [84,85].

Intriguingly, certain predictions made by the aberrant precision hypothesis can be formally tested via deploying predictive processing modeling. The latter approach allows for the tracking of potentially critical processes of the hypothesized “predictive brain” and may, therefore, have the potential to become an invaluable tool for revisiting the condition of autism. To date, a number of different theoretical and computational predictive processing and active inference models have been put forward, covering a variety of levels, functions and temporal scales. In the next section, we will suggest modeling examples of potential relevance to the autism research on the individual level. More specifically, we will view here predictive processing and active inference as a common framework for re-addressing traditional ideas about ASC. The “five big ideas”, which rest on diverse functions of a cognitive level, will motivate and help to structure our suggestion.

#### 4. Individual level: Predictive processing and active inference as a common framework for integrating diverse neurocognitive hypotheses

Theory of Mind (ToM) - as described above - can be viewed as an inference problem [86], where the brain tries to understand ‘invisible’ mental states, through observable human behavior. Koster-Hale and Saxe review evidence that relates ToM to predictive processing formulations [87]. To that end, they consider how relevant brain regions such as the superior temporal sulcus (STS), temporoparietal junction (TPJ) and medial prefrontal cortex (MPFC) might be involved in mental state inference across different time scales. To be more specific, STS has been implicated in neural reactions to face and body action in the scale of seconds, while TPJ has been related to assessing desires and beliefs of other people, which can last from minutes to years and MPFC has been thought to contribute to the evaluation of temporally more stable traits of other people.

The social motivation (SM) hypothesis of autism focuses on how a lack of motivation for processing and learning about social aspects could be relevant for understanding ASC or how difficulties in social cognition could decrease interest for social cues. Interestingly, Heyes has argued that social learning shares the same basic cognitive mechanisms with non-social learning [88]. In line with this, Behrens *et al.* indicated that standard reward-based associative processes guide the acquisition of social information as well [89]. More specifically, they showed activation of the anterior cingulate cortex gyrus (ACCs) and sulcus (ACCg) for reward-based and social learning respectively. At the level of decision-making, it was found that ventromedial prefrontal cortex (VMPFC) encodes both probabilities about the social and non-social sources, appearing to integrate information from ACCs and ACCg in a subject-specific fashion (see also [90,91]). Consequently, the above mentioned brain regions could potentially play an important role in the investigation of ASC related differences in multimodal cue integration and contextualization of precision in social and non-social cues.

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

**Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press**

As previously discussed, the so-called “mirror neuron system” has also been implicated in ASC via the broken mirror neuron (BMN) hypothesis. According to the BMN hypothesis, difficulties in ASC in understanding others’ actions and intentions may arise from a defective functioning of the mirror neuron system. However, precisely how mirror neurons contribute to action/intention understanding is still unclear [92]. Kilner and colleagues suggested that the brain deploys a mirror neuron predictive processing model and minimizes prediction error at all levels [93]. More specifically, they considered a hierarchy that consists of four levels of decreasing abstraction descending the hierarchy; the (1) intention, (2) goal, (3) kinematic, and (4) muscle levels respectively (see [94]). These levels of behavior are generally assumed to be independent of each other [94]. This assumption, however, appears not to be true as recent evidence indicates that the kinematics of a performed movement already reflect the agent’s intention and makes it distinguishable [95]. This raises the intriguing possibility that intentions may be decoded from movement kinematics [96]. A reasonable framework for integrating different sources of prediction is that a range of possible intentions is first estimated from the spatial and temporal context, e.g. in predictive areas outside the mirror system [92]. This prior prediction can impact on action understanding, constraining the number of possible intentions. Early movement discriminant kinematic features of the observed motor act, can lead then to the selection of the most probable intention. Studying such inference problems in light of predictive processing and active inference could provide further insights on the implications of a broken mirror neuron account for understanding ASC.

Visual processing and particularly the extraction of spatiotemporal regularities might also be related to specific theories about ASC, such as the weak central coherence hypothesis (WCC). Natural images tend to be correlated both in space and time. That is, natural scenes usually consist of finite regions of relatively uniform attributes and tend to reflect region-specific uniform intensity values [97]. For example, a stable object, being viewed from a constant perspective appears to emit relatively similar intensity values over time. These regular spatio-temporal characteristics can be exploited by the visual system to predict intensity values in advance based on neighboring and historical information. Indeed, Rao and Ballard proposed that the brain predicts this kind of regularities via a predictive processing model [98], embodied in neural loops of increasing receptive fields as ascending the hierarchy (e.g. the lateral geniculate nucleus (LGN) – primary visual cortex (V1) – secondary visual cortex (V2) feedback loop [97]). Such a family of models could be exploited in the future for an investigation of aspects related to a weak central coherence in ASC and more precisely the extraction of perceptual regularities. For instance, quantifying autism-specific styles in extracting such regularities could yield further insights about facts as perceptual hypersensitivity and differences in perceiving certain kind of illusions (e.g. see [99]).

The executive dysfunction (ED) hypothesis focuses on executive cognition and behavior. Kopp has recently emphasized the relevance of executive function for predictive processing theories [100]. More precisely, drawing on the latter and self-terminating operating units [101], Kopp proposed a theoretical hierarchical model for dealing with executive dysfunction, especially focusing on brain regions as the medial, orbital and lateral prefrontal cortex (PFC). Indeed, there is evidence speaking for a hierarchical

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

***Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press***

organization of the rostro-caudal axis of the PFC, based on the level of abstraction (see [102,103]). We suggest such kind of models could prove to be fruitful in studying putative executive dysfunction through the hierarchical inference entailed by predictive processing and active inference in ASC.

Lawson and colleagues have recently put forward several suggestions with regard to potentially aberrant predictive processing processes relevant for an understanding of ASC at a neurobiological level as well [60]. For instance, plasma oxytocin, which has been suggested to control the relative salience of social and non-social stimuli [41], has been found to be reduced in children with ASC [38]. These can be linked to an aberrant precision hypothesis, under the assumption that oxytocin is involved in contextualizing precision of social as compared non-social stimuli (see [104]).

Taken together, we suggest that a multitude of aspects in ASC can be integrated under the predictive processing and active inference perspective. *By doing so, ASC can be revisited as a different prediction and (inter-)action style, as opposed to a set of a priori impaired neurocognitive functions that reside in specific brain regions.* This exact shift of perspective, however, begs the question of how does such a different style emerge? In the next section we tackle this question, by leaning on sociocultural historical theories, which emphasize the social construction of the (a-)typical self and Bayesian accounts of brain function, which provide a powerful toolbox for the investigation of underlying mechanisms.

## 5. Integrating individual and collective levels of analysis: The dialectical misattunement hypothesis

We open this section by discussing different approaches which although following distinct lines of argument converge on the idea that focusing on individual brains will not be enough to fully understand the human mind and psychopathology. In particular, we will argue against considering only biological mechanisms, since in our view, the latter reductionist approach covers only part of the dialectical interplay between individual processes and the collective level of analysis. In fact, cultural historical activity theories have strongly emphasized the importance of considering the interrelationship between individual and socio-cultural processes in psychological and psychopathological research: For instance, Vygotsky already distinguished social interaction as a key factor in the formation of consciousness and ‘higher’ human psychological processes, which he argued are developed *through and due to social interactions* [6]. Additionally, he claimed that every function appears twice in a child’s development, first on a social level (i.e. “intermind”) and then on an individual level (i.e. “intramind”): *“All the higher functions originate as actual relationships between individuals”* [5]. In other words, he suggested that through communication, through the direct social interaction with others a child *internalizes* active cultural values in society (as cited in [6]), realizing that the (a-)typical self is dialectically and socially constructed.

Interestingly, recent developments in accounts of social cognition and intersubjectivity have also focused on the enabling or even constitutive role of social interaction [15,16,95,105–124]. More specifically, mainstream accounts of social cognition have been criticised for neglecting the interactive dimension of social situations and for adopting an individualistic view of (social) cognition (e.g.

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

***Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press***

specifically on the example of autism see [125], philosophical considerations [107] and neuroscientific research [120]). With regard to psychiatric conditions it has also been suggested that transdiagnostically observed social impairments are more likely or may only manifest under conditions of real-time social interaction, whereas situations of social observation might be less problematic [123]. Furthermore, several accounts have been critical toward core assumptions of contemporary cognitivist paradigms, which have been thought of as viewing the brain or more generally the organism merely as a passive 'consumer' of external stimuli [126]. Despite each account's distinct commitments, these kinds of approaches are usually positioned under the umbrella of the 4Es [127,128], which described cognition as enactive [129–131], embodied [132–134], embedded [130,132,133], extended [130,135], but also affective [136,137]. In line with these accounts, using scenarios of higher ecological validity, which do not neglect the critical role of the body, the environment and interactions in cognition, could offer a more suitable framework to study brain function and behavior [16,120].

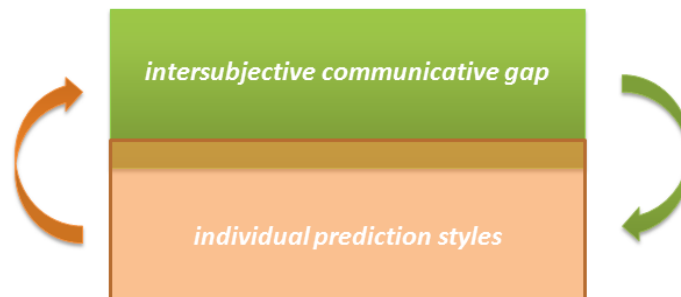
On top of providing a naturalistic scenario, interactive situations also potentially allow for the consideration of turn-taking (e.g. [112]) and emergent social phenomena at higher levels of description, which otherwise might remain intangible (e.g. [15]). In neuroscience, cognition has generally tended to imply a dynamic interaction between brain areas merely within a single skull. However, there is no theoretical reason to *a priori* exclude other body parts, and generally other people, as well as mediating cultural tools, as cultural historical activity theories would emphasize. In line with an enactivist or dynamical systems perspective two or more communicating agents can be seen as a coupled system, being driven by non-linear interactions (e.g. [113,114,138]). However, investigating individual predictive processing mechanisms in order to understand communicative processes between agents could also be particularly informative. Notably, a formal account of addressing communication as reciprocal exchange of predictions about the other's behavior has recently been put forward [139,140]: This account, which rests on predictive processing, takes into consideration both perceptual updating and action expression within a closed loop between two agents. Here, simulations were used to illustrate how two agents, which model each other, could in theory converge into a system of generalized synchrony (i.e. synchronization of chaos), thereby effectively embodying a single shared model. In contrast to this 'solipsistic' understanding of communication, we argue that by adopting a dialectical perspective we will look for such synchronization dynamics across different levels of description and do not assume that my understanding of another is realized entirely in my own head.

To be more specific, we suggest that a '*dialectical misattunement*' constitutes one of the defining factors of ASC and other psychiatric conditions. We suggest that communication misalignments and weak interpersonal coupling in social interactions might be the result of increasingly divergent predictive and (inter-)action styles across individuals (cf. predictive processing and active inference). From an ontogenetic perspective, such a misattunement could result in impoverished opportunities for acquiring socio-culturally mediated knowledge and skills. *In other words, we view two potentially cardinal processes that are tightly intertwined in a dialectical relationship: at the collective level weak coupling, crucially modulated by sociocultural factors, might lead to greater inter-individual incompatibilities in prediction and (inter-)action styles, while at the individual level, diverging prediction and (inter-)action*

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

***Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press***

styles might lead to weak communicative coupling with others in social interaction.<sup>5</sup> In short, ‘*dialectical misattunement*’ refers to an imbalance between individual and collective levels, rather than exclusively considering single levels. This view particularly highlights the critical role of social interaction into human development and the social construction of the (a-)typical self. Consequently, the interactive nature of social situations can help to enhance or decrease differences in prediction and (inter-)action style in a feedback-loop fashion (cf. the circular causality introduced above). That is, small initial differences in the individual level are thought of cumulatively enhancing (or weakening) interpersonal coupling during social interaction and vice versa. Schematically, an initial communicative gap, could yield incompatible prediction and action styles and vice versa (**Fig. 3**).



**Figure 3.** Dialectical Misattunement: increasing communicative gap (collective level) yields increasingly different prediction and action styles (individual level) and vice versa.

Notably, such communicative misattunement could be expected to unfold across *multiple temporal scales*; for example this could take place during the course of a dialogue (scale of minutes), during a human relationship (scale of months or years), or along development (scale of a lifetime). Additionally, with regard to groups of people (e.g. the so called psychopathological groups, or generally any other social group), this kind of misattunement could even take on a cultural form, spanning a scale of several generations. For instance, culturally cultivated beliefs in a given society about a specific group of people (e.g. stereotypes) might modulate the communication efficacy between in- and out-group persons.

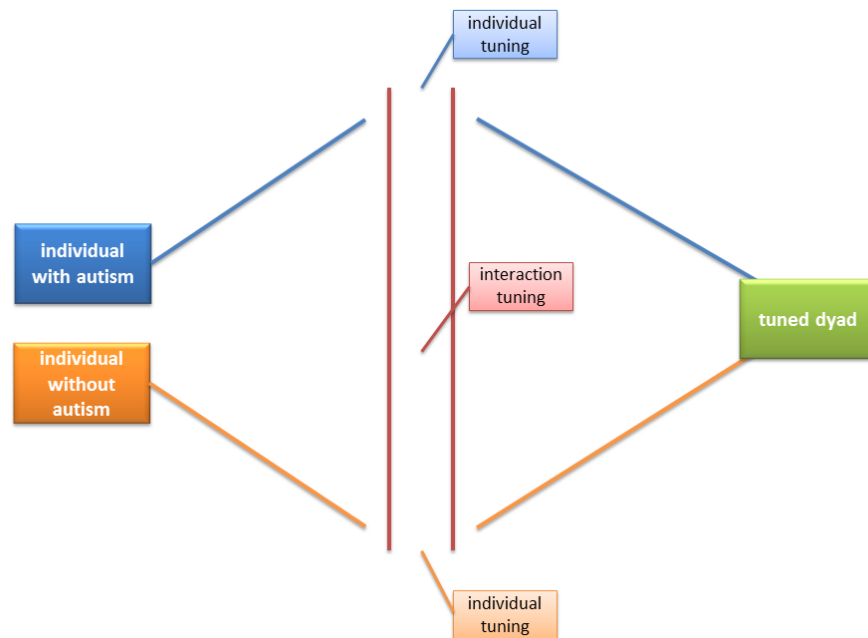
More broadly, we believe that for gaining a complete understanding of conditions such as ASC, a shift of focus from the individual brain to the interaction between people, is essential. Intriguingly, as we will argue in the next and final section, such an approach could yield formal insights into both individual and collective mechanisms (cf. [15]), as well as intra- and inter-condition communication characteristics. Additionally, in psychiatry it could facilitate research at both a diagnostic and a treatment level. In short, we view the future of relevant theoretical research and clinical practice not only as an investigation of ‘*disordered*’ brain mechanisms, but of a ‘*misattunement*’ between persons as well. In line with the dialectical misattunement of social interaction hypothesis, which highlights intersubjectivity as an indispensable factor of human development, we also suggest the enrichment of approaches which

<sup>5</sup> Please note misattunement encompasses both aspects of dissimilarity (e.g. social misalignment) and non-complementarity (e.g. dysregulated coupling).



exclusively aim at ‘tuning’ the person with the ASC. To this end, we suggest considering tuning also the ‘other’ (i.e. the neurotypical person who the person with ASC interacts with), as well as the social interaction medium (i.e. sociocultural framework, such as social expectations and stereotypes, as well as technological medium such as educational social robotics) [16].

More precisely, in a clinical setting, one could, therefore, pay attention not only to the potentially “maladaptive” processes within the diseased individual, but to the coupling dynamics of the dyad (for instance during psychotherapy or group sessions) and critically the interaction between the individual and the collective. Additionally, our approach also motivates an alternative pedagogical program. The latter would primarily aim at tuning, not merely individual behavior, but crucially the interaction between people. Here, the pedagogical procedure would move beyond the traditional classroom, focusing on cognitive and behavioral aspects of not only the person with a specific condition (e.g. ASC), but also their interactors (e.g. parents, educators, or other peers) and most importantly communication and mediating factors (**Fig. 4**).



**Figure 4.** Schematic presentation of a misattunement amelioration: by intervening both on the individual (e.g. cognitive and behavioral training of both interactors) and the collective level (e.g. adjustments of cultural/technological tools, sociofeedback). (blue: individual trajectory of a person with ASC; orange: individual trajectory of a person without ASC; trajectories here represent multiple temporal scales, from minutes in the course of a conversation to years across development).

This could be achieved by developing adjustable frameworks, both to the individual and the interaction itself. A promising solution could be found in the form of ‘smart’ technology, which could track and guide traditional educational practice, taking into account real-time activity, but also historically relevant

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

**Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press**

aspects (e.g. [141]). Crucially, while biofeedback techniques have been fruitfully used for monitoring and constructively exploiting individual activity (e.g. physiological factors), our approach would further point toward an extended notion of feedback, here referred to as ‘*sociofeedback*’, including relational parameters as well (e.g. interpersonal coupling). Furthermore, the proposed shift of attention could be beneficial not only in clinical and pedagogical practice, but more broadly with regard to societal practice.

For instance, by diffusing ideas in society about viewing psychiatric conditions, as disorders of social interaction rather than disorders of individuals [123], psychiatric stigma could be attenuated. As Vygotsky used to highlight, simply speaking, aspects of specific difficulties related to psychiatric conditions can be thought of as falling into two main categories; first aspects which are directly related to a biological level, and second aspects which are related to relevant beliefs and practices in society. Although social processes play a decisive role in shaping a person’s mental reality, emphasis is usually only given to biology. Notably, such difficulties being a social product to a large extent, they could be historically (along both social-historical and individual-developmental trajectories) alleviated. Furthermore, our approach emphasizes the dialectical relation of the collective and the individual (e.g. interrelations between culture and individual persons, as in interactions between ‘patient’ and ‘examiner’, or ‘patient’ and ‘non-patient’). The broadened scope of effective treatment could encompass both personal and interpersonal parameters. In this light, the relativity of psychiatric diagnosis, which is usually the outcome of a communicative procedure between a potential patient and a culturally tuned examiner (e.g. psychiatrist or psychologist), also becomes more evident (cf. [10,16,115,123,142]). In technical terms, our approach could be reframed as studying potential dynamic and recurrent feedback loops across and within different levels of description, as well as temporal scales, driving both quantitative and qualitative changes (cf. dialectics). We believe that computational modeling, such as Bayesian accounts, as well as dynamical systems approaches can prove to be fruitful tools for scientifically testing the potentials of such a perspective. In fact, in our closing section we will motivate a *Bayesian account of intersubjectivity*, which will aim at formally accommodating both individual and collective mechanisms.

## 6. Summary and Outlook: From a synthesis of dialectical and computational approaches to a Bayesian account of intersubjectivity

In this article, taking dialectics as a point of departure and drawing upon insights from multiple areas of research we have argued that considering inherent interrelations as well as integrating findings from diverse levels of description, within-level processes and multiple temporal scales will be essential in future autism research. Such a holistic development, we claim, will help to unveil the intrinsic units of analysis for reconstructing the critical dimensions of a multilevel and multidimensional condition such as ASC; thus here thought of as an ‘*autism space*’, rather than a *spectrum*. In particular, we discussed how a framework such as the predictive processing and active inference could be used for bringing together and re-addressing under a common umbrella, traditional hypotheses at the level of the individual (e.g. neurobiology, cognition and behavior). By doing so, ASC was revisited as a different prediction and (inter-)action style, as opposed to a set of *a priori* impaired neurocognitive functions that reside in

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

***Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press***

specific brain regions. Then we argued that such an approach is not sufficient on its own, but needs to be directed towards the relevant real-life phenomena that take place during social interaction. Consequently, we propose an approach for integrating a computational and a dialectical perspective to psychiatric conditions for scientifically studying both intra- and inter-personal processes by introducing the ‘*dialectical misattunement*’ hypothesis of social interaction. Misattunement across persons is thought of as disturbances of the dynamic and reciprocal unfolding of an interaction across multiple time scales, resulting in increasingly divergent prediction and (inter-)action styles (ways of generating and expressing expectations about the (social) world and the self). This thesis does not consider psychiatric conditions, such as ASC, merely as disordered function within individual brains, but rather as an interactive mismatch between persons.

In a forthcoming paper we will use the conceptual arguments introduced above to illustrate the dialectical misattunement hypothesis formally. Specifically, we will analyze two-person simulations and experiments [16] with dual hierarchical Gaussian filters (HGF; [143]), as a formal (computational) model of dyadic exchange [15]. This provides a quantitative and principled description of the dialectical misattunement hypothesis - and how it could be verified empirically using relatively simple paradigms and analyses. In concrete terms, we suggest that established techniques of multilevel computational modeling (e.g. [143,144]) can be used to investigate the interrelation of individual brain mechanisms and interpersonal processes. Intrasubjective parameters (e.g. on the dynamics of belief updating) will be used for modeling individual brain processes of two (or more) brains, while intersubjective will be introduced on a second meta-Bayesian level for capturing dyadic (or group collective) processes, such as interpersonal coupling [15]. The latter scheme will thus move beyond current neuromodeling approaches by also considering emergent phenomena on higher levels of description, such as questions about the autonomy of a dyad or a group of people and the individuality of the mind. To give a more specific example, in the context of collective decision-making or joint action, a non-linear model might optimally explain observed behavior, thus, providing evidence that the dyad or the group is different than the sum of individuals. Inversely, this framework could address questions about how mechanisms of societal structure, and in general collective processes, in turn, shape individual reality. For instance, one could differentially study the potentially distinct impact, which a competitive versus a collaborative structure might exert upon an individual. Notably, this kind of modeling architecture will not be merely able to model multiple levels of description, but interlevel processes as well (e.g. internalization and externalization mechanisms).

Moving the focus from the observation of *individual observers*, toward a multilevel observation of *dyads and groups of interactors* could help to explore whether and how interpersonal coordination might actually serve as a prior and modulate the need for inferences about hidden causes of social behavior. Such an intersubjectively Bayesian approach, we claim, will provide a formal characterization of subject-specific, as well as dyad- and group-level dynamics. It will, thereby, significantly advance our understanding of ASC and other psychiatric conditions thought of as disorders of social interaction. As we provocatively state in the title of this article we suggest we need to go beyond autism, not by neglecting the existence of the condition, but by adopting a holistic approach, which will embrace the

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

**Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press**

individual with autism, as well as the socioculturally mediated interactions with other people. The ultimate goal of such an approach will be to go beyond current diagnostic and treatment practice by promoting a reciprocal alignment of individual and societal practices as opposed to a single-sided adjustment of individual behavior and brain function into the ‘normal’.

**Table 1.** Glossary of terms, as they either appear in the bibliography (e.g. Stanford Encyclopedia of Philosophy) or were introduced in this article.

| Table 1                                      | Glossary of terms   |
|--|---|
| <b>Active inference</b>                      | An account of action according to which (biological) systems sample the environment in accordance with prior beliefs, for minimizing free energy.   |
| <b>Bayesian account of intersubjectivity</b> | The ‘Bayesian account of intersubjectivity’ is considered here as a Bayesian account of human activity, that takes into account both intrapersonal and interpersonal processes.   |
| <b>Bayesian inference</b>                    | A method of statistical inference in which Bayes' theorem is used to update the probability for a hypothesis as more evidence or information becomes available.   |
| <b>Biofeedback</b>                           | A training technique by which a person learns how to regulate certain body functions, such as heart rate, blood pressure, or brain wave patterns, that are normally considered to be involuntary.   |
| <b>Cultural-historical psychology</b>        | Theory aiming at accounting for the inseparable unity of mind, brain and culture, in their development (and/or degradation) in concrete socio-historical settings (in case of individuals) and throughout the history of humankind (as socio-biological species). |
| <b>Circular causality</b>                    | Common situation in complex systems (with several interconnected causes and effects) where an action is controlled or affected by its own outcome or result.  |

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

***Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press***

|                                  |   |
|----------------------------------|---|
| <b>Dialectics</b>                | The dialectical method states that phenomena can be understood only in their wholeness, inner contradiction and movement.   |
| <b>Dialectical misattunement</b> | The ‘dialectical misattunement hypothesis’ rethinks a psychiatric condition such as ASC, not merely as a disorder of the individual brain, but as cumulative misattunement between persons as well. Misattunement across persons can be thought of as disturbances of the dynamic and reciprocal unfolding of an interaction across multiple time scales, resulting in increasingly divergent prediction and action styles. |
| <b>Dualism</b>                   | Theory stating that for some particular domain, there are two fundamental kinds or categories of things or principles (e.g. the physical and the mental).   |
| <b>Enactivism</b>                | Philosophical account that argues that cognition arises through a dynamic interaction between an acting organism and its environment.   |
| <b>Emergence</b>                 | Phenomenon whereby larger entities arise through interactions among smaller or simpler entities such that the larger entities exhibit properties the smaller/simpler entities do not exhibit.   |
| <b>Free energy</b>               | Principle that tries to explain how (biological) systems maintain their order by restricting themselves to a limited number of states.  |
| <b>Heterogeneous dyads</b>       | ‘Heterogeneous dyads’ are considered here dyads consisted of either only neuro-typical persons, or only persons with a certain condition, as ASC.   |
| <b>Homogeneous dyads</b>         | ‘Homogeneous dyads’ are considered here dyads consisted of persons with different conditions, such as one neurotypical person and one person with ASC.  |
| <b>Interaction tuning</b>        | Interaction tuning here refers to tuning of expectations of either or both interactors, as well as facilitating a social interaction via tuning the communication medium.   |
| <b>Intrapersonal</b>             | (Processes) being unfolded within the person.   |

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

***Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press***

|                                     |   |
|-------------------------------------|---|
| <b>Interpersonal</b>                | (Processes) being unfolded between persons.   |
| <b>Precision</b>                    | A statistical term defined as the inverse variance and can be thought of as the confidence a (biological) system places upon its beliefs.   |
| <b>Prediction error</b>             | The discrepancy between incoming information and a (biological) system's generated predictions.   |
| <b>Predictive coding/processing</b> | Theory that states that (biological) systems are constantly generating and updating hypotheses that predict sensory input at varying levels of abstraction, for minimizing free energy.   |
| <b>Reductionism</b>                 | A belief that the whole of reality consists of a minimal number of parts.   |
| <b>Sociofeedback</b>                | ' <i>Sociofeedback</i> ' is considered here a (future) training technique by which a person, a dyad or a group of people will learn how to (co-)regulate certain social interaction processes, such as interpersonal coupling and coordination. The concept also applies to automatic adjustment of the interaction medium, based on social interaction monitoring. |
| <b>Synergetics</b>                  | An interdisciplinary science explaining the formation and self-organization of patterns and structures in open systems far from thermodynamic equilibrium.  |

## Acknowledgments

Dimitris Bolis would like to express his gratitude to colleagues in Translational Neuromodeling Unit (ETH Zurich), Autism Research Group (Trinity College Dublin), and the Independent Max Planck Research Group for Social Neuroscience (MPI of Psychiatry). In particular, Dimitris is grateful to Klaas Enno Stephan and Karl Friston, as well as two anonymous reviewers for helpful feedback, but would also like to warmly thank Kyveli Kompatsiari, Andreas Miliadis, Falk Lieder, Codruta Sudrijan, Angeliki Dotsi and Leontios Hadjileontiadis, for inspiring discussions on a multitude of topics, from Vygotsky and Bayes to enactivism and autism.

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

**Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press**

## References

- 1 Wong W: Understanding Dialectical Thinking from a Cultural-Historical Perspective. *Philos Psychol* 2006;19:239–260.
- 2 Dafermos M: Rethinking the Crisis in Social Psychology: A Dialectical Perspective. *Soc Personal Psychol Compass* 2015;9:394–405.
- 3 Bidell T: Vygotsky, Piaget and the Dialectic of Development. *Hum Dev* 1988;31:329–348.
- 4 Lawler J: Dialectical Philosophy and Developmental Psychology: Hegel and Piaget on Contradiction. *Hum Dev* 1975;18:1–17.
- 5 Vygotsky LS (1980).: *Mind in society: The development of higher psychological processes*. Harvard University Press.
- 6 Vagenas E: *The Soviet psychology for the development of the child*. Synchroni Epohi, 2003.
- 7 Hoel EP, Albantakis L, Tononi G: Quantifying causal emergence shows that macro can beat micro. *Proc Natl Acad Sci* 2013;110:19790–19795.
- 8 Haken H: *Synergetics : introduction and advanced topics : with 266 figures*
- 9 Vygotsky LS: *Thought and language* 1964;14:97–98.
- 10 Dafermos M: *The cultural historical theory of Vygotsky*. Atrapos, 2002.
- 11 Stephan KE, Mathys C: Computational approaches to psychiatry. *Curr Opin Neurobiol* 2014;25:85–92.
- 12 Huys QJM, Maia T V, Frank MJ: Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat Neurosci* 2016;19:404–413.
- 13 Friston K: The free-energy principle: a unified brain theory? *Nat Rev Neurosci* 2010;11:127–138.
- 14 Bolis D, Becchio C, Schilbach L: Revisiting Psychological Definitions at the Interface of Sociocultural Historical Theories and Predictive Coding; in : *Welcome to WORLDING THE BRAIN: Patterns, Rhythms, Narratives in Neuroscience and the Humanities*. 2016, p 10.
- 15 Bolis D, Schilbach L: Beyond one Bayesian brain: Modeling intra- and inter-personal processes during social interaction: Commentary on “Mentalizing homeostasis: The social origins of interoceptive inference” by Fotopoulou & Tsakiris. *Neuropsychanalysis* 2017;19:35–38.
- 16 Bolis D, Schilbach L: Observing and participating in social interactions: Action perception and action control across the autistic spectrum. *Dev Cogn Neurosci* 2017; DOI: 10.1016/j.dcn.2017.01.009
- 17 Wing L: The History of Ideas on Autism. *Autism* 1997;1:13–23.
- 18 Dakin S, Frith U: Vagaries of Visual Perception in Autism. *Neuron* 2005;48:497–507.

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

***Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press***

- 19 Plaisted Grant K, Davis G: Perception and apperception in autism: rejecting the inverse assumption. *Philos Trans R Soc B Biol Sci* 2009;364:1393–1398.
- 20 Treffert DA: The savant syndrome: an extraordinary condition. A synopsis: past, present, future. *Philos Trans R Soc London B Biol Sci* 2009;364.
- 21 Frith U, Happé F: Autism: beyond “theory of mind”. *Cognition*50:115–132.
- 22 Baron-Cohen S: Is asperger syndrome/high-functioning autism necessarily a disability? *Dev Psychopathol* 2000;12:489–500.
- 23 Frith U: *Autism : a very short introduction*. Oxford University Press, 2008.
- 24 Baron-Cohen S, Leslie AM, Frith U: Does the autistic child have a “theory of mind” ? *Cognition* 1985;21:37–46.
- 25 Senju A, Southgate V, White S, Frith U: Mindblind Eyes: An Absence of Spontaneous Theory of Mind in Asperger Syndrome. *Science* (80- ) 2009;325.
- 26 Schilbach L, Eickhoff SB, Cieslik EC, Kuzmanovic B, Vogeley K: Shall we do this together? Social gaze influences action control in a comparison group, but not in individuals with high-functioning autism. *Autism* 2012;16:151–162.
- 27 di Pellegrino G, Fadiga L, Fogassi L, Gallese V, Rizzolatti G: Understanding motor events: a neurophysiological study. *Exp Brain Res* 1992;91:176–180.
- 28 Rizzolatti G, Craighero L: THE MIRROR-NEURON SYSTEM. *Annu Rev Neurosci* 2004;27:169–192.
- 29 Oberman LM, Hubbard EM, McCleery JP, Altschuler EL, Ramachandran VS, Pineda JA: EEG evidence for mirror neuron dysfunction in autism spectrum disorders. *Cogn Brain Res* 2005;24:190–198.
- 30 Ramachandran VS, Oberman LM: Broken Mirrors: A Theory of Autism. *Sci Am* 2006;295:62–69.
- 31 Dapretto M, Davies MS, Pfeifer JH, Scott AA, Sigman M, Bookheimer SY, et al.: Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nat Neurosci* 2006;9:28–30.
- 32 Perkins T, Stokes M, McGillivray J, Bittar R: Mirror neuron dysfunction in autism spectrum disorders. *J Clin Neurosci* 2010;17:1239–1243.
- 33 Southgate V, de C. Hamilton AF: Unbroken mirrors: challenging a theory of Autism. *Trends Cogn Sci* 2008;12:225–229.
- 34 Fan Y-T, Decety J, Yang C-Y, Liu J-L, Cheng Y: Unbroken mirror neurons in autism spectrum disorders. *J Child Psychol Psychiatry* 2010;51:981–988.
- 35 Wang Y, Hamilton AF de C: Social top-down response modulation (STORM): a model of the control of mimicry in social interaction. *Front Hum Neurosci* 2012;6:153.
- 36 Dumas G, Soussignan R, Hugueville L, Martinerie J, Nadel J: Revisiting mu suppression in autism

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

**Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press**



- spectrum disorder. *Brain Res* 2014;1585:108–119.
- 37 Chevallier C, Kohls G, Troiani V, Brodtkin ES, Schultz RT: The social motivation theory of autism. *Trends Cogn Sci* 2012;16:231–239.
- 38 Modahl C, Green L, Fein D, Morris M, Waterhouse L, Feinstein C, et al.: Plasma oxytocin levels in autistic children. *Biol Psychiatry* 1998;43:270–7.
- 39 Bartz JA, Zaki J, Bolger N, Ochsner KN: Social effects of oxytocin in humans: context and person matter. *Trends Cogn Sci* 2011;15:301–309.
- 40 Modi ME, Young LJ: The oxytocin system in drug discovery for autism: Animal models and novel therapeutic strategies. *Horm Behav* 2012;61:340–350.
- 41 Gordon I, Vander Wyk BC, Bennett RH, Cordeaux C, Lucas M V, Eilbott JA, et al.: Oxytocin enhances brain function in children with autism. *Proc Natl Acad Sci* 2013;110:20953–20958.
- 42 Frith U: *Autism : explaining the enigma*. Blackwell Pub, 2003.
- 43 Happé F: Autism: cognitive deficit or cognitive style? *Trends Cogn Sci* 1999;3:216–222.
- 44 Happé F, Frith U: The Weak Coherence Account: Detail-focused Cognitive Style in Autism Spectrum Disorders. *J Autism Dev Disord* 2006;36:5–25.
- 45 Mottron L, Dawson M, Soulières I, Hubert B, Burack J: Enhanced Perceptual Functioning in Autism: An Update, and Eight Principles of Autistic Perception. *J Autism Dev Disord* 2006;36:27–43.
- 46 Murray D, Lesser M, Lawson W: Attention, monotropism and the diagnostic criteria for autism. *Autism* 2005;9:139–156.
- 47 Steel MJG, Gorman R, Flexman JE: Neuropsychiatric Testing in an Autistic Mathematical Idiot-Savant: Evidence for Nonverbal Abstract Capacity. *J Am Acad Child Psychiatry* 1984;23:704–707.
- 48 Rumsey JM: Conceptual problem-solving in highly verbal, nonretarded autistic men. *J Autism Dev Disord* 1985;15:23–36.
- 49 Ozonoff S, Pennington BF, Rogers SJ: Executive Function Deficits in High-Functioning Autistic Individuals: Relationship to Theory of Mind. *J Child Psychol Psychiatry* 1991;32:1081–1105.
- 50 Hill EL: Executive dysfunction in autism. *Trends Cogn Sci* 2004;8:26–32.
- 51 Happé F: Cognition in Autism: One Deficit or Many?; in : *Neural Basis and Treatment Possibilities: Novartis Foundation Symposium 251*. John Wiley & Sons, Ltd, 2003, pp 198–212.
- 52 Happé F, Ronald A, Plomin R: Time to give up on a single explanation for autism. *Nat Neurosci* 2006;9:1218–1220.
- 53 Gallagher S, Varga S: Conceptual issues in autism spectrum disorders. *Curr Opin Psychiatry* 2015;28:1.

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

***Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press***

- 54 Pellicano E, Burr D: When the world becomes “too real”: a Bayesian explanation of autistic perception. *Trends Cogn Sci* 2012;16:504–510.
- 55 Brock J: Alternative Bayesian accounts of autistic perception: Comment on Pellicano and Burr. *Trends Cogn Sci* 2012;16:573–574.
- 56 Friston KJ, Lawson R, Frith CD: On hyperpriors and hypopriors: comment on Pellicano and Burr. *Trends Cogn Sci* 2013;17:1.
- 57 van Boxtel JJA, Lu H: A predictive coding perspective on autism spectrum disorders. *Front Psychol* 2013;4:19.
- 58 Van de Cruys S, De-Wit L, Evers K, Boets B, Wagemans J: Weak Priors versus Overfitting of Predictions in Autism: Reply to Pellicano and Burr ( TICS , 2012). *Iperception* 2013;4:95–97.
- 59 Teufel C, Subramaniam N, Fletcher PC: The role of priors in Bayesian models of perception. *Front Comput Neurosci* 2013;7:25.
- 60 Lawson RP, Rees G, Friston KJ: An aberrant precision account of autism. *Front Hum Neurosci* 2014;8:302.
- 61 Van de Cruys S, Evers K, Van der Hallen R, Van Eylen L, Boets B, De-Wit L, et al.: Precise minds in uncertain worlds: Predictive coding in autism. *Psychol Rev* 2014;121:649–675.
- 62 Sevgi M, Diaconescu AO, Tittgemeyer M, Schilbach L: Social Bayes: Using Bayesian Modeling to Study Autistic Trait–Related Differences in Social Cognition. *Biol Psychiatry* 2016;80:112–119.
- 63 Knill DC, Pouget A: The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci* 2004;27:712–719.
- 64 Harris H, Israeli D, Minschew N, Bonnef Y, Heeger DJ, Behrmann M, et al.: Perceptual learning in autism: over-specificity and possible remedies. *Nat Neurosci* 2015;18:1574–1576.
- 65 Mumford D: On the computational architecture of the neocortex. *Biol Cybern* 1992;66:241–251.
- 66 Friston K: A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci* 2005;360:815–836.
- 67 Friston K: Hierarchical Models in the Brain. *PLoS Comput Biol* 2008;4:e1000211.
- 68 Clark A: Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav Brain Sci* 2013;36:181–204.
- 69 Gomot M, Wicker B: A challenging, unpredictable world for people with Autism Spectrum Disorder. *Int J Psychophysiol* 2012;83:240–247.
- 70 Helmholtz HLF Von: *Handbuch der physiologischen Optik* 1867; DOI: 10.3931/E-RARA-21259
- 71 William James: *The Principles Of Psychology*. Read Books Ltd.
- 72 Stock A, Stock C: A short history of ideo-motor action. *Psychol Res* 2004;68:176–188.
- 73 Herbot O, Butz M V.: Too Good to be True? Ideomotor Theory from a Computational  
*This is a post-print version of the accepted manuscript. Until official publication please cite as:*  
**Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press**

- Perspective. *Front Psychol* 2012;3:494.
- 74 Felleman DJ, Van Essen DC: Distributed Hierarchical Processing in the Primate Cerebral Cortex
- 75 Friston K, Kiebel S: Predictive coding under the free-energy principle. *Philos Trans R Soc B Biol Sci* 2009;364:1211–1221.
- 76 Friston KJ, Stephan KE: Free-energy and the brain. *Synthese* 2007;159:417–458.
- 77 Friston KJ, Daunizeau J, Kilner J, Kiebel SJ: Action and behavior: a free-energy formulation. *Biol Cybern* 2010;102:227–260.
- 78 Clark A: The many faces of precision (Replies to commentaries on &quot;Whatever next? Neural prediction, situated agents, and the future of cognitive science&quot;). *Front Psychol* 2013;4:270.
- 79 Feldman H, Friston KJ: Attention, Uncertainty, and Free-Energy. *Front Hum Neurosci* 2010;4. DOI: 10.3389/fnhum.2010.00215
- 80 Shipp S, Adams RA, Friston KJ: Reflections on agranular architecture: predictive coding in the motor cortex. *Trends Neurosci* 2013;36:706–716.
- 81 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.
- 82 Friston K: The free-energy principle: a rough guide to the brain? *Trends Cogn Sci* 2009;13:293–301.
- 83 Hohwy J: The predictive mind
- 84 Van de Cruys S, Wagemans J: Putting Reward in Art: A Tentative Prediction Error Account of Visual Art. *Iperception* 2011;2:1035–1062.
- 85 Joffily M, Coricelli G, FitzGerald T, Galea JM, Adams R: Emotional Valence and the Free-Energy Principle. *PLoS Comput Biol* 2013;9:e1003094.
- 86 Baker CL, Saxe RR, Tenenbaum JB: Bayesian Theory of Mind: Modeling Joint Belief-Desire Attribution
- 87 Koster-Hale J, Saxe R: Theory of Mind: A Neural Prediction Problem. *Neuron* 2013;79:836–848.
- 88 Heyes C: What’s social about social learning? *J Comp Psychol* 2012;126:193–202.
- 89 Behrens TEJ, Hunt LT, Woolrich MW, Rushworth MFS: Associative learning of social value. *Nature* 2008;456:245–249.
- 90 Apps MAJ, Green R, Ramnani N: Reinforcement learning signals in the anterior cingulate cortex code for others’ false beliefs. *Neuroimage* 2013;64:1–9.
- 91 Balsters JH, Apps MAJ, Bolis D, Lehner R, Gallagher L, Wenderoth N: Disrupted prediction errors index social deficits in autism spectrum disorder. *Brain* 2017;140:235–246.
- 92 Kilner JM: More than one pathway to action understanding. *Trends Cogn Sci* 2011;15:352–357.

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

**Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press**

- 93 Kilner JM, Friston KJ, Frith CD: Predictive coding: an account of the mirror neuron system. *Cogn Process* 2007;8:159–166.
- 94 Hamilton AF de C, Grafton ST: The motor hierarchy: from kinematics to goals and intentions; in : *Sensorimotor Foundations of Higher Cognition*. Oxford University Press, 2007, pp 381–407.
- 95 Becchio C, Manera V, Sartori L, Cavallo A, Castiello U: Grasping intentions: from thought experiments to empirical evidence. *Front Hum Neurosci* 2012;6:117.
- 96 Ansuini C, Cavallo A, Bertone C, Becchio C: Intentions in the brain: the unveiling of Mister Hyde. *Neuroscientist* 2015;21:126–35.
- 97 Huang Y, Rao RPN: Predictive coding. *Wiley Interdiscip Rev Cogn Sci* 2011;2:580–593.
- 98 Rao RPN, Ballard DH: Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* 1999;2:79–87.
- 99 Happé FG: Studying weak central coherence at low levels: children with autism do not succumb to visual illusions. A research note. *J Child Psychol Psychiatry* 1996;37:873–877.
- 100 Kopp B: A simple hypothesis of executive function. *Front Hum Neurosci* 2012;6:159.
- 101 Miller GA: *Plans and the structure of behavior*. Holt, 1960.
- 102 Badre D, D’Esposito M: Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *J Cogn Neurosci* 2007;19:2082–2099.
- 103 Badre D, D’Esposito M: Is the rostro-caudal axis of the frontal lobe hierarchical? *Nat Rev Neurosci* 2009;10:659–69.
- 104 Quattrocki E, Friston K: Autism, oxytocin and interoception. *Neurosci Biobehav Rev* 2014;47:410–430.
- 105 Gallagher S: The practice of mind: Theory, simulation or primary interaction? *J Conscious Stud* 2001;8:83–108.
- 106 Schilbach L, Wohlschlaeger AM, Kraemer NC, Newen A, Shah NJ, Fink GR, et al.: Being with virtual others: Neural correlates of social interaction. *Neuropsychologia* 2006;44:718–730.
- 107 De Jaegher H, Di Paolo E: Participatory sense-making. *Phenomenol Cogn Sci* 2007;6:485–507.
- 108 Oullier O, de Guzman GC, Jantzen KJ, Lagarde J, Scott Kelso JA: Social coordination dynamics: Measuring human bonding. *Soc Neurosci* 2008;3:178–192.
- 109 Reddy V: *How infants know minds*. Harvard University Press, 2010.
- 110 Hari R, Kujala M V: Brain Basis of Human Social Interaction: From Concepts to Brain Imaging. *Physiol Rev* 2009;89:453–479.
- 111 De Jaegher H, Di Paolo E, Gallagher S: Can social interaction constitute social cognition? *Trends Cogn Sci* 2010;14:441–447.

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

**Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press**

- 112 Schilbach L, Wilms M, Eickhoff SB, Romanzetti S, Tepest R, Bente G, et al.: Minds Made for Sharing: Initiating Joint Attention Recruits Reward-related Neurocircuitry. *J Cogn Neurosci* 2010;22:2702–2715.
- 113 Froese T, Di Paolo EA: The enactive approach Theoretical sketches from cell to society\*. *Pragmat {&} Cogn* 2011;19:1–36.
- 114 Froese T, Fuchs T: The extended body: a case study in the neurophenomenology of social interaction. *Phenomenol Cogn Sci* 2012;11:205–235.
- 115 Di Paolo E, De Jaegher H: The interactive brain hypothesis. *Front Hum Neurosci* 2012;6:163.
- 116 Konvalinka I, Roepstorff A: The two-brain approach: how can mutually interacting brains teach us something about social interaction? *Front Hum Neurosci* 2012;6:215.
- 117 Dumas G: Towards a two-body neuroscience. *Commun {&} Integr Biol* 2011;4:349–352.
- 118 Dumas G, Martinerie J, Soussignan R, Nadel J: Does the brain know who is at the origin of what in an imitative interaction? *Front Hum Neurosci* 2012;6:128.
- 119 Timmermans B, Schilbach L, Pasquali A, Cleeremans A: Higher order thoughts in action: consciousness as an unconscious re-description process. *Philos Trans R Soc London B Biol Sci* 2012;367.
- 120 Schilbach L, Timmermans B, Reddy V, Costall A, Bente G, Schlicht T, et al.: Toward a second-person neuroscience. *Behav Brain Sci* 2013;36:393–414.
- 121 Dale R, Fusaroli R, Duran ND, Richardson DC: The Self-Organization of Human Interaction; in : *Psychology of Learning and Motivation - Advances in Research and Theory*. Academic Press Inc., 2013, pp 43–95.
- 122 Schilbach L: On the relationship of online and offline social cognition. *Front Hum Neurosci* 2014;8:278.
- 123 Schilbach L: Towards a second-person neuropsychiatry. *Philos Trans R Soc London B Biol Sci* 2015;371.
- 124 Hyperscanning neuroimaging technique to reveal the “two-in-one” system in social interactions. *Neurosci Res* 2015;90:25–32.
- 125 Gallagher S: Understanding Interpersonal Problems in Autism: Interaction Theory as An Alternative to Theory of Mind. *Philos Psychiatry, {&} Psychol* 2004;11:199–217.
- 126 Tom Froese: Overview | Dr. Tom Froese [cited 2017 Sep 9]; Available from: <https://froese.wordpress.com/research/overview/>
- 127 Menary R: Introduction to the special issue on 4E cognition. *Phenomenol Cogn Sci* 2010;9:459–463.
- 128 Ward D, Stapleton M: Es are good 2012, pp 89–104.

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

**Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press**

- 129 Varela FJ, Thompson E, Rosch E: The embodied mind : cognitive science and human experience. MIT Press, 1991.
- 130 Hurley SL: Consciousness in action. Harvard University Press, 1998.
- 131 Thompson E: Mind in life : biology, phenomenology, and the sciences of mind. Belknap Press of Harvard University Press, 2007.
- 132 Haugeland J: Having thought : essays in the metaphysics of mind. Harvard University Press, 1998.
- 133 Clark A: Being there : putting brain, body, and world together again. MIT Press, 1997.
- 134 Gallagher S: How the Body Shapes the Mind. Oxford University Press, 2005. DOI: 10.1093/0199271941.001.0001
- 135 Clark A, Chalmers D: The Extended Mind. *Analysis* 1998;58:7–19.
- 136 Colombetti G: Enactive appraisal. *Phenomenol Cogn Sci* 2007;6:527–546.
- 137 Ratcliffe M: Existential feeling and psychopathology 2009;
- 138 Dumas G, de Guzman GC, Tognoli E, Kelso JAS: The human dynamic clamp as a paradigm for social interaction. *Proc Natl Acad Sci* 2014;111:E3726--E3734.
- 139 Friston K, Frith C: A Duet for one. *Conscious Cogn* 2015;36:390–405.
- 140 Friston KJ, Frith CD: Active inference, communication and hermeneutics. *Cortex* 2015;68:129–143.
- 141 Vittorias J, Petrantonakis P, Bolis D, Tsiligkyri A, Kosmidou V, Hadjileontiadis LJ: NOESIS: An Enhanced Educational Environment for Kids with Autism Spectrum Disorders; in : 2008 Eighth IEEE International Conference on Advanced Learning Technologies. IEEE, 2008, pp 1019–1020.
- 142 Laing RD, Phillipson H, Lee AR: Interpersonal perception; a theory and a method of research. Harper & Row, 1972.
- 143 Mathys C, Daunizeau J, Friston KJ, Stephan KE: A bayesian foundation for individual learning under uncertainty. *Front Hum Neurosci* 2011;5:39.
- 144 Bolis D, Heinze J, Mathys C, Stephan KE: Inferring cognitive traits of individual subjects through gaze controlled video games 2015
- 145 Pessoa L: On the relationship between emotion and cognition. *Nat Rev Neurosci* 2008;9:148–158.

*This is a post-print version of the accepted manuscript. Until official publication please cite as:*

***Bolis et al. (in press) Beyond autism: Introducing the dialectical misattunement hypothesis and a Bayesian account of intersubjectivity. ©S. Karger AG, Psychopathology; in press***