Abstract

In the present work, we explore the development of processing of emotional facial configurations under a Predictive Processing (or Predictive Coding) framework. Predictive Processing provides a new approach to brain function that has been used to explain a wide range of processes, from perception to socio-emotional processing. The explanatory power of this framework for adult brain function is widely recognized, but it has yet to be systematically applied to understanding the developing brain. Studying the findings of developmental research under this framework may allow a deeper understanding of the predictive mechanisms and their ontogenetic course, and adds to knowledge on brain functions and developmental processes. Therefore, the goal of this work was to explore the potential complementarity of predictive processing and development. Specifically, we focus on how the development of facial and emotion processing may be understood under a predictive processing framework. The processing of facial expressions was selected due to the developmental relevance of these stimuli, their impact on general emotional development, as well as the large body of literature on this topic (comprised of both well-established but also incongruent findings, which a novel approach may clarify). Considering the main findings of developmental research on the processing of emotion-related facial configurations under this framework, we argue that predictive processing is consistent with developmental evidence and provides a promising avenue for developmental research, as it reveals new questions in the fields of development and emotion processing.

Keywords: predictive processing; predictive coding; development; face processing; emotion processing; facial expressions of emotion

Understanding the Development of Face and Emotion processing under a Predictive Processing Framework

Although there is no consensus on how to define development, there are common aspects shared by all conceptions of development, such as the occurrence of dynamic changes throughout the lifespan, the role of plasticity, and the intricate interplay between organisms and their environments at multiple levels of analysis (Lerner, Hershberg, Hilliard, & Johnson, 2015). Simultaneously, Predictive Processing (PP, or Predictive Coding) approaches are growing in popularity as a broad explanatory framework for how neural and psychological systems work and change over time (Clark, 2013). Indeed, both developmental and PP approaches are concerned with how experience drives internal neural and psychological mechanisms that in turn shape future experiences. However, the necessary efforts made to link the two fields are still emerging in the literature.

The brain has been theorized as a somewhat passive system that generates complex representations over time, through the accumulation of evidence resulting mostly from bottom-up stimulation, with top-down processes playing a modulatory role. However, in the last decade, the emerging framework of Predictive Processing has challenged this conception, viewing the brain as an active organ that is constantly predicting its future state and stimulation (Clark, 2013; Friston, 2005, 2010). This framework has grown in popularity due to its neurobiological plausibility (e.g., Barrett & Simmons, 2015; Bastos et al., 2012; Deneve, Alemi, & Bourdoukan, 2017) and great explanatory power for several low- and high-level processes, leading to a more embodied perspective of cognition and emotion (Clark, 2015; Varela, Thompson, & Rosch, 1993). In the present work, we are approaching emotional development from a PP perspective, specifically regarding the processing of facial expressions.

The scientific interest in human faces and facial “expressions” is due to their special perceptual significance when compared to other visual inputs starting from early in development (e.g., Goren, Sarty, & Wu, 1975; Johnson, Dziurawiec, Ellis, & Morton, 1991). Face processing is also intriguing due to the existence of brain structures and networks that seem to be specialized for this process (Cohen-Kadosh & Johnson, 2007; Johnson, 2011). The human face is among the most relevant cues in social interactions. It allows collecting information about the person’s identity, sex, and age, and the dynamic aspects of facial movement and configuration (e.g., configuration of facial muscles, eye gaze, mouth movements) provide cues about psychological states (Bruce & Young, 1986). Facial communication plays a preponderant role in pre-verbal infants as an instrument to communicate their needs, to establish fruitful attachment relationships, and to understand the world through the reactions of others (de Haan & Nelson, 1997; Grossmann, 2010).

For example, face recognition in the developing child allows the recognition of the primary caregivers that address the child’s needs. The effect of this familiarity can be observed in the reunion episode of the Strange Situation, in which the recognition of the caregiver, together with other caregiver’s behaviors, contributes to the regulation of the child’s affect (e.g., Ainsworth & Bell, 1970). Additionally, the process of extrinsic emotional regulation by the caregiver is dependent upon adequate facial communication (Gross, 2013; Gross & Thompson, 2007). This is thought to involve the child’s ability to understand the emotional content of the caregiver’s facial configurations and to use that information for emotional regulation. This has traditionally been studied under the term facial “expressions” of emotion, in which the configurations of facial muscles are thought to directly reflect an internal emotional state. This concept will be critically discussed below (specifically distinguishing between facial configurations and the emotional interpretation of that configuration), but, for the present examples, we refer to facial “expressions” as traditionally interpreted. For instance, in the above example, a positive facial “expression” by the caregiver during the reunion can be soothing and pleasant for the child and help the child better regulate her emotional state. facial “expressions” are also important in situations involving social referencing in which the child looks to the caregiver’s facial “expression” to reduce the ambiguity of the situation and to behave accordingly (Feinman, 1992). For example, negative expressions by caregivers may signal dangerous situations and children are likely to interrupt their behavior.

As already mentioned, the term facial “expression” of emotion, or just facial “expression”, implies that certain configurations of facial muscles are manifestations of internal emotional states and that such facial configurations lead to the perception of the associated emotions. For early theories of Basic Emotions, these two processes (facial configuration and internal emotional states) were expected to go hand-in-hand: each emotion (such as fear or happiness) would consist of a set of innate, species-universal components that include facial expressions of emotion (Ekman & Cordaro, 2011). From this perspective, the role of development would be to simply allow the components of emotion to come on-line as the child matures, since they are hypothesized to form a cohesive set, shaped by evolution (e.g., Izard, 1977; as critiqued in Barrett, 2013).

Developing the ability to understand emotional states from facial configurations requires both the visual processing of the facial configurative features (facial processing) and the awareness and interpretation of the emotional state of the other person (emotion-related processing). Indeed, there is developmental evidence that suggests that understanding the emotional meaning from facial configurations is significantly shaped by experience (e.g., Pollak, Messner, Kistler, & Cohn, 2009) and follows a protracted developmental pathway, possibly beginning with a broad comprehension based on valence (pleasant vs. unpleasant) that only progressively becomes more defined to allow distinguishing between specific emotional categories (Widen, 2013). This latter conception of emotional development is more consistent with the definition of development advanced in the beginning of this paper and with the emerging PP frameworks for neural and psychological functioning.

Therefore, our goal was to examine how the development of a “predictive mind” (as phrased by Hohwy, 2013) may account for the development of the ability to process facial configurations and the gradual link between these and what adults refer to as “emotional categories”. In order to do so we start by briefly introducing the PP model. Then, we synthesize the literature on the processing of facial configurations during development. We specifically consider how perceptual discrimination of facial configurations (facial processing) and the understanding of the emotional meaning of those displays (emotion-related processing) may be conceptualized within a single predictive hierarchy. Finally, we discuss the future implications of the PP framework for understanding emotional development and the developing brain.

A brief introduction to Predictive Processing

We have selected "predictive processing" as an umbrella term to capture the general idea that neural and cognitive systems are best described under a predictive architecture; that is, as systems that process information by continuously generating predictions and comparing them to the actual inputs, as described below (see Gallagher & Allen, 2016, for a more nuanced proposal). We will discuss how PP has been used to account for perception in terms of Perceptual Inference (perhaps the most well-known aspect of PP approaches) and supplement this with a depiction of action under a PP framework, namely by considering Active Inference. It should be noted that the PP framework refers to principles about neural and psychological function, rather than the ability to make conscious predictions in everyday life. An alternative way of framing this approach would be to refer to it as Bayesian Inference (this point will be further developed below), regarded as the general computational goal of integrating prior knowledge with incoming sensory information (Aichison & Langyel, 2017).

Perceptual Inference

Although PP is a contemporary framework in the context of brain function, the main concept was described in the late nineteenth century, by Helmholtz (1866/1962; Clark, 2013; Friston, 2010), as "unconscious inference”. Helmholtz explained that, since retinal inputs are ambiguous, previous knowledge is required for perception in order to give sense to, and infer some properties of, the visual object. In fact, the first field in which predictive coding was implemented as an algorithm was visual processing (e.g., Rao & Ballard, 1999; Spratling, 2017). Current formulations suggest that the brain is an active organ, constantly making predictions about its own future states or stimulation (Clark, 2013). The rationale is that the brain, in a certain context, builds a prediction and compares it with the actual sensory input. When there is a match, this means that the sensory input was successfully predicted and no updating of the internal prediction is required. However, if there is a mismatch, then the prediction failed to account for the incoming sensory input and the difference between prediction and input generates a Prediction Error (PE). PEs can be regarded as the information that remains to be explained in the input and serves to update subsequent predictions associated with the input (Mumford, 1992). By updating predictions through the incorporation of PEs, the system will minimize PE in the next round of comparisons, thus creating a better (predictive) model of the input. This process is thought to occur hierarchically throughout the cortex in a sequence of Prediction-PE loops, possibly implemented in cortical microcircuits along sequentially connected cortical regions (e.g., Barrett & Simmons, 2015; Bastos et al., 2012).

At the first level, predictions are about immediate sensory input, but at the next level of the hierarchy, predictions are about the predictions of the lower level. Furthermore, PEs at each level of the hierarchy reflect the mismatch between the prediction from the higher level and activity from the lower level. In this hierarchy, it is then possible to predict not only the immediate input (yielding rapid perceptual inference) but also more stable regularities of the environment (yielding slower timescale perceptual learning; Clark, 2015), in what may be considered a generative model of the causal structure of the world (Allen & Friston, 2016; Clark, 2013; Hohwy, 2013). This means that, early in the predictive hierarchy, there will be predictions and prediction errors about very fast occurring and local characteristics of the stimuli, but that, further along the hierarchy, predictions and prediction errors will, in turn, refer to long-lasting features of the stimuli and context. A general illustration of this scheme is depicted in Figure 1.

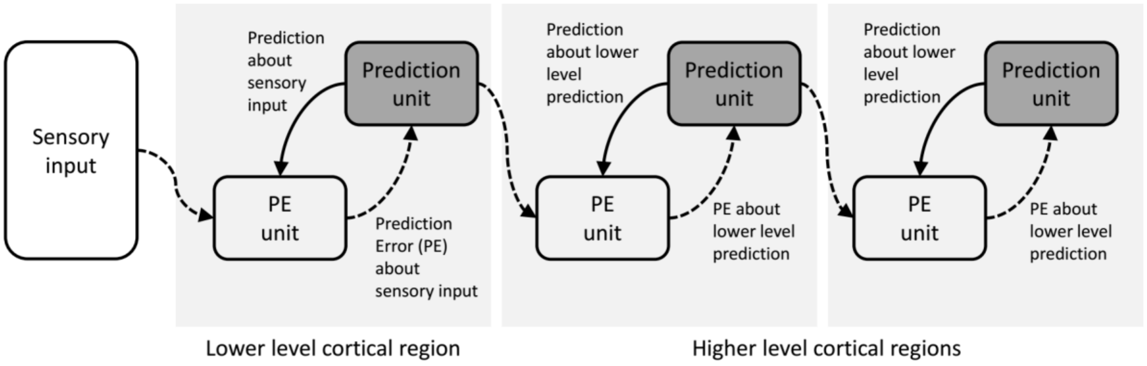

Figure 1. Simplified scheme of a general Predictive Processing (PP) model of the cortical hierarchy for Perceptual Inference (for much more in-depth computational and anatomical treatments of PP see, e.g., Barrett & Simmons, 2015; Bastos et al., 2012; Friston, 2005, 2010).

Numerous algorithms have been proposed to represent the implementation of predictive coding. As depicted in a recent review by Spratling (2017), these algorithms vary in aspects such as the processes believed to be optimal for the integration of the input in the existing model, or the neurobiological underpinnings of these processes. Moreover, not all predictive coding algorithms implicate Bayesian inference computations (reviewed in Atchinson & Lengyel, 2017), as will be discussed below. However, all predictive processing models consider that (a) the main goal is to fit a model of the world to a specific input, thus relying on previous information that is generated by prior knowledge, and (b) the accommodation of the new input leads to a more accurate posterior internal model of the external world (Friston, 2012; Spratling, 2017).

As previously mentioned, the PP model provides an alternative to traditional conceptions of the brain as a somewhat passive system that mainly elaborates on bottom-up stimulation with some degree of top-down modulation. The bottom-up theories of information processing in the brain postulate that processing typically begins with the passive reception of information (an input or stimulus) that is sequentially fed forward and processed throughout cortical regions, possibly leading to a behavioral outcome (Spruit, 2008). Although this has been applied to a wide variety of processes, some phenomena remained unexplained. Nonetheless, it is important to emphasize that the PP model does not reject the existence of bottom-up processing. Instead, it considers both directions of information flow as necessary. The (feed-forward) PEs will be used to update an unconfirmed prediction or will be suppressed by a correct top-down (feedback) prediction. It is also the case that the precision of the bottom-up signal will influence the confidence in the sensory input. In statistics, the precision of a parameter estimate is defined as the inverse of the standard deviation of the estimator of that parameter (Everitt & Skrondal, 2010). In Bayesian modeling, the larger the precision of the evidence (i.e., the smaller the variance of the likelihood distribution), the more it will influence the posterior distribution. Conversely, noisy or unstable sensory signals (high variability/low precision) will influence the posterior distribution to a lesser extent. In terms of PP, the statistical precision of sensory inputs and

PEs refers to whether these signals are reliable enough to update the prediction. The term "precision-weighted PEs" captures this notion by highlighting that the ability of PEs to update internal predictions is dependent on a weight that is specified by how statistically precise the signals are (for a more detailed account of the role of precision in PP see: Ferreira-Santos, 2016; Friston, 2008; Kok & De Lange, 2015).

Active Inference

So far, we have discussed a PP approach to perception, defined as the interplay between predictions and PEs across the cortical hierarchy. In Perceptual Inference, PEs are minimized by updating internal predictions, resulting in perceptual learning of an internal model that represents external stimuli. However, this is still a passive view of perceptual learning. Given that PEs are computed as the mismatch between prediction and sensory input, PEs can also be minimized without modifying the internal prediction by altering the sensory input itself, namely through action (Friston, Mattout, & Kilner, 2011). The concept of Active Inference suggests it may occur in the service of Perceptual Inference (Friston et al., 2016). By acting on the environment, we may test how our actions affect our sensory input and verify whether those sensory inputs deviate from our internal predictions. This is important because, in many cases, the sensory input that is passively available may not be sufficient to infer what the stimulus is. For a rudimentary but illustrative example, we may see a person's face from a lateral angle and suspect this person is crying, but cannot be certain from that viewing angle. By moving in front of the person, we may reduce that uncertainty by either confirming the initial prediction of a crying face if the person was indeed crying, or by updating that prediction based on the PE if the person was not. This is made possible as corollary discharges or efferent copies of activity from the motor system may enable sensory systems to predict the sensory outcomes of actions (Clark, 2013, 2015).

Another consequence of the general principle of Active Inference, namely the modification of PEs by action rather than by updating the perceptual model, is that it can support goal-directed behavior. Indeed, moving in and acting on the environment may produce changes to the sensory input that lead to a decrease in PE, as the novel input gradually conforms to an internal model representing a goal state (Friston et al., 2016). In this scenario, the predictive model is kept unchanged and the action is used to change the environment to produce the desired sensory consequences (Hohwy, 2013; Shadmehr, Smith, & Krakauer, 2010). For instance, if a child intends to eat a candy and believes there is one inside a closed box, she may hold an internal model of herself perceptually experiencing finding and eating the candy, while actively opening the box and acting to reduce PE to that intended situation.

In the brief accounts put forth above, we have considered Perceptual Inference and Active Inference separately for expository purposes, but it should be noted that sensory and motor functions are inherently coupled. Indeed, one of the strengths of predictive functioning in perception is the adaptive advantage it confers. The ability to form predictions about the incoming states allows the brain to deal with the uncertainty present in the world (Knill & Pouget, 2004) and to allostatically prepare for future states accordingly (Sterling, 2012). Moreover, it represents an economic use of neural resources by reducing redundancy and avoiding the processing of the whole stimulus each time it is perceived (Huang & Rao, 2011). This provides a suitable answer to the bottleneck issues associated with the previous theories of neural information processing. By processing only a reduced portion of the signal, the system reduces the risk of information loss that is inherent to this limited-capacity system (Clark, 2015). We would like to stress at this point that some computational and anatomical details of the predictive processing models are simplified for brevity (for a more detailed treatment of these aspects see, e.g., Aichison & Langyel, 2017; Barrett & Simmons, 2015; Friston, 2005, 2010; Spratling, 2017).

Predictive Processing and Development

By appealing to a parsimonious fundamental mechanism (e.g., Bastos et al., 2012) and integrating a multiplicity of processes at different levels of analysis (Friston, 2008), PP provides a comprehensive and integrative explanatory model for the understanding of both low- and high-level neurocognitive processes. However, PP has not been systematically applied to understanding the developing brain, as it has been for the adult brain. Indeed, a predictive neural and psychological system is a system that changes over time as a function of the experience it accrues and integrates in persisting predictive models, which, in turn, shape the following experiences. This definition is strikingly similar to a general definition of development as a dynamical process of increasing neural and psychological complexity through interactions with the physical and social environment, and shaping future interactions based on the development of those neural and psychological structures. To further highlight the links between PP and development we will discuss the emergence of internal representations from a PP perspective as a model for developmental change.

Building mental representations involves both exposure to the stimuli to be represented and a system of pattern detection that allows for feature analysis. All models of cognition agree on this process but they differ in their concept of representation and how much it influences perception (e.g., Firestone & Scholl, 2016). The PP literature often equates predictions with probabilistic representations of the world (Gładziejewski, 2016). These probabilistic representations are considered to be generated by “representation units”, and are defined as a prototype of the structure of a given phenomenon. In turn, this prototypical representation is thought to guide our perceptual, motor, and cognitive behavior towards the phenomenon. As Gładziejewski (2016) states, representations are “action-guiding, detachable, structural models that afford representational error detection" (p. 569), and the idea of predictions as embodied signals is also mentioned by Barrett (2017b). In addition to "representation" and "prior", the internal model is also sometimes termed "belief" and is defined as the "probability distribution over some unknown state or attribute of the external world" (Adams, Brown, & Friston, 2014, p.52). This term is often complemented by the consideration of the prior as the belief that is present before the event, and the posterior as the updated belief after the experience.

Regarding the computational implementations, PP models do not necessarily implement Bayesian inference computations. However, there are findings that support the combination of both (a) predictive coding algorithms in behavior and brain function and (b) Bayesian inference computations in neural processing (Aitchison & Lengyel, 2017; Lochmann & Deneve, 2011). Thus, assuming a Bayesian stance, a representation is a prior belief about some future state. Hence, we need to understand the Bayes rule for prediction. The Bayes rule is commonly defined as a way of updating our belief (B) about some hypothesis given new evidence (E), allowing inferences to be made based on uncertain information using probabilities. The posterior belief P(B|E) is the conditional probability of our belief given the evidence and may be computed by the multiplication of our prior belief P(B) by the likelihood that the evidence occurs given that our belief is correct P(E|B). By labeling B as internal model and E as sensory input, the Bayes rule follows the scheme depicted in Figure 2. For more details, please consult the simplified simulated example of Bayesian modeling on the topic of the present review that we provide as Supplemental Material.

$$\overbrace{P(\mathit{internal\ model}|\mathit{sensory\ input})}^{\text{Posterior}} = \frac{\overbrace{P(\mathit{sensory\ input}|\mathit{internal\ model})}^{\text{Likelihood}}\overbrace{P(\mathit{internal\ model})}^{\text{Prior}}}{P(\mathit{sensory\ input})}$$

Figure 2. Bayes rule for perception, considering the prior as the internal model and the evidence as the sensory input.

Prior beliefs can be learned and updated by experiencing the world, thus contributing to our knowledge of the probabilistic features of the phenomena. Very early in development, there is no previous inherent extensive knowledge and, therefore, no extensive internal model of the world. At that point, there is a constant updating of the prior as the first contacts with sensory stimuli generate an abundance of prediction errors. As initial PEs begin to be accommodated the consequent representations become stronger and develop into better predictions, or beliefs, with more accurate probabilistic distributions. Clark (2013) describes three-stages of belief building: (a) a lack of commitment to a single interpretation and abundance of errors; (b) a convergence to a dominant general theme; and (c) a progressive specification of representational details for both the percept and associated contextual features.

The idea of predictive processes operating in early stages of development is not new. Currently, there is evidence of predictive processing mechanisms in infants (Trainor, 2012), as well as relevant discussions about the fundamental role of these mechanisms in the development of learning abilities (Trainor, 2012), interoception, and allostasis (Atzil & Barrett, 2017).

In recent years, there have been important advances in PP studies of face processing in adults, suggesting that face-sensitive cortical regions are responsive to the predictability of facial stimuli (Brodski, et al., 2015) and seem to pre-activate prior knowledge in expectation of such stimuli (Apps & Tsakiris, 2013; Brodski-Guerniero, et al. 2017; Trapp, Schweinberger, Hayward, & Kovács, 2018). Our expectation is that, by extending the PP approach to the development of processing of what have been typically called facial “expressions” of emotion (i.e., facial configurations that may be interpreted in terms of emotional meaning), we may be able to provide an example of how to integrate PP and development in a way that is likely to generate new research avenues, especially for the field of emotional development.

Development of face perception and processing of emotion-related facial configurations

Development of face processing and PP

The special significance of the face when compared to the rest of the visual world is already observable in early stages of development. Beyond the orientation to faces observed in newborns (e.g., Farroni et al., 2013; Johnson, Dziurawiec, Ellis, & Morton, 1991), recent studies uncovered the presence of orientation to face-like stimuli in fetal samples (Braddick, 2017; Reid et al., 2017). Additionally, 3-month-old infants show different neural activity for faces with direct, compared to averted, gaze (Farroni, Csibra, Simion, & Johnson, 2002), which may confer social relevance to these stimuli beyond a mere perceptual preference. A model of neurocognitive development that has been proposed for understanding brain specialization, including face perception, is Johnson’s model of Interactive Specialization (IS; Johnson, 2001, 2005b, 2011). In this model, the cortical and cognitive specialization that is observed during development results from interactions between the environment and the brain, and between different structures within the brain. One chief example of IS is the development of specialized neural circuits for face perception. In this framework, the preferences that newborns show for orienting to faces (e.g., Goren, Sarty, & Wu, 1975) is explained by the existence of a visual bias towards faces (or face-like objects) that enhances and ensures a continued perception of facial stimuli. In turn, this continued exposure to facial stimuli drives the development of the regions of the visual cortex, tuning it for face processing and effectively resulting in perceptual expertise for face perception. Therefore, functional cortical regions that initially are poorly defined are repeatedly stimulated with a class of inputs and specialize in a certain group of stimuli, such as upright human faces. This specialization is reflected in a narrower functioning of these circuits, which begin to respond more selectively and faster to face stimuli (de Haan, Johnson, & Halit, 2007; Johnson, 2001, 2005b, 2011).

The assumptions of the previous model converge with some essential features and hypotheses of PP, particularly the early bias towards faces and the role of perceptual experience in this specialization. Given the evidence of visual processing *in utero* (Reid et al., 2017), it is possible that the face prior in newborns derives from pre-natal visual statistical learning. This prior may be a very primitive conception of a face, such as a visually top-heavy stimulus with a face-like configuration - as seen in the infants' preference for face-like patterns as simple as three blobs located in the place of the eyes and the mouth or faces with their inner features scrambled but which maintain the top-heavy distribution (Cassia, Turati, & Simion, 2004; Farroni et al., 2005; Johnson, 2005; Simion, Leo, Turati, Valenza, & Dalla Barba, 2007). The establishment of a face-like prior may, subsequently, be the basis for the early tendency of newborns to look at and visually follow faces. This active and voluntary sampling of the environment possibly illustrates an early form of Active Inference. That is, in contact with the visual world, the infant searches for the predicted face-like stimuli in order to fulfil this prediction and actively reduce the PEs coming from the sampling of the rest of the environment. On the other hand, the experience of the visual world, rich in facial stimuli, and the accommodation of the forthcoming prediction errors leads to the specialization in this category of stimuli via processes of Perceptual Inference. In addition, by considering the existence of circuits and structures that are computationally ready to process these stimuli, we can assume there is an inherent expectation of the existence and relevance of faces, which is akin to what has been described as experience-expectant development (Greenough, Black, & Wallace, 1987). Finally, if face stimuli are actively sampled this could explain the fact that face representations are developed earlier than any other visual stimulus (de Haan, Johnson, & Halit, 2007). In sum, the mechanisms that are the basis of face specialization, according to Johnson (2011), are largely consistent with the presence of PP from early developmental stages.

Moreover, more recent models also consider this interaction between brain and environment from a predictive processing viewpoint. For instance, Barrett (2017b) describes the occurrence of constant attempts to predict and give sense to the world from early developmental stages, so as to build interoceptive predictions and models of the world, which are shaped by experience and, in turn, influence future experiences.

This bias towards faces and the consequent specialization of the visual system is hypothesized to lead to the development of a “face space” during the first months of life (Simion & Di Giorgio, 2015; Slater et al., 2010). This space is defined as a continuum consisting of all the faces that have been experienced (or by the set of experienced face characteristics) in which the average face occupies the center and that becomes narrower around the center with increasing exposure to face stimuli (Nelson, 2001, 2003; O'Toole, Castillo, Parde, Hill, & Chellappa, 2018; Valentine, 1991). To some extent, this concept is consistent with PP, if one considers the face space as the probabilistic space of possible priors centered around the average representation of facial stimuli determined by previous experience (for a more detailed predictive coding approach to perceptual narrowing and face space, see Balas, 2012). The capacity for face categorization and the development of the “face space” seems to be present in infants by around 3-month of age (Simion & Di Giorgio, 2015). By then, infants show the ability to distinguish between different facial identities and to detect an average face among a group of observed faces, which will be perceived as being more familiar and, therefore, elicit less processing and attention than each individual face (de Haan, Johnson, Maurer, & Perrett, 2001). This average face, which may serve as a prior, has been shown to depend on the experiences of the child, as by 3- and 4-months of age there is a preference for faces of the same gender of the caregiver (Quinn, Yahr, Kuhn, Slater, & Pascalis, 2002), and for faces of one's own ethnic group (Bar-Haim, Ziv, Lamy, & Hodes, 2006; Gaither, Pauker & Johnson, 2012; Kelly et al., 2007). Furthermore, the experiences in these first months that determine the statistical properties of the face prior are likely related to the neural specialization for processing face stimuli, contributing to this specialization and also being influenced by it (discussed in Smith, Jayaraman, Clerkin, & Yu, 2018). The presence of face-processing expertise and sensitivity to more detailed configuration changes seem to be present by 5-months of age, concomitant with the improvement of visual acuity (Diamond & Carey, 1986; Hayden, Bhatt, Reed, Corbly, & Joseph, 2007; Leder & Bruce, 2000) and cumulative exposure to face stimuli, as documented by video monitoring with infant head cameras (Jayaraman, Fausey, & Smith, 2017).

In summary, there is a well-established evidence base for a gradual cortical specialization for face processing in early development that is consistent with PP accounts. Whether this developmental process is integral or tangential to processing facial configurations is still a matter of debate (Calder, 2011), but the argument may serve as a hypothesis for the development of the ability to perceptually discriminate facial configurations in infants.

Development of emotion-related processing of facial configurations and PP

Research on the development of the abilities (a) to process facial configurations which adults typically consider to be exemplars of emotional categories and (b) to infer emotional states has traditionally been reported using the term facial "expression" of emotion to refer to both processes without explicitly identifying which process is being considered. As stated above, we will refer to facial processing as the ability to decode and recognize different facial configurations, and to emotion-related processing as the inference of affective or emotional states that may rely on facial configurations (but potentially include other sources of information). We use the term facial “expressions” of emotion (in quotation marks) when referring to the way the term was originally expressed in the literature. Many studies focused on one or many specific emotions, each presumed to be "expressed" by a corresponding facial configuration. We will refer to these facial configurations by reporting in italics the "emotional" category that was intended by the authors (e.g., *happy*).

Taking this into account, we will now analyze the developmental course of the processing of emotion-related facial configurations, targeting how does it transitions from perceptual discrimination to the detection and understanding of emotional content.

Perceptual discrimination of emotion-related configurations

Regarding processing of emotion-related facial configurations, 3-month-old infants seem to be able to discriminate between some different facial configurations (de Haan & Matheson, 2009). Until 7-months of age, this discrimination is known to be mainly between *happiness* and other "emotional" categories (e.g., Barrera & Maurer, 1981; Young-Browne et al., 1978) and between different intensities of the same "emotional" category, such as in *happiness*, *fear*, and *anger* (for a review see Grossmann, 2010). Nonetheless, as pointed out by Grossmann (2010), in order for the expressions to convey emotional meaning and content to the infants, they must be able to understand that each facial configuration expresses the same emotional meaning, regardless of the individual who performs it. In other words, that the children must be able to dissociate between changes in "emotional" content and changes in identity, and to recognize that the "emotional" category itself remains the same despite changes in intensity of the facial configuration. Otherwise, the discrimination will only be dependent on visual perceptual changes.

However, differences in facial configurations, even if recognized across different individuals, may still be achieved merely by facial processing. Indeed, the fact that a child can recognize the same facial configuration in different individuals (e.g., recognize that two different adults share the same expression because they both show low eyebrows) does not mean that she is able to understand the emotional meaning (e.g., understand that they may be experiencing *anger* and the meaning of that affective state). As discussed in the introduction, this intrinsic knowledge about emotions and emotion processing is postulated by Basic Emotions Models but it fails to capture the phenomena and processes that seem to characterize emotional development. For example, a study by Kaneshige and Haryu (2014) reported that 4-month-old infants distinguished between facial configurations of *anger* and *happiness* but reacted positively to both, which may imply a lack of knowledge about the affective meaning of each facial configuration at this age. This means that although the infant may be able to differentiate between some different facial configurations, additional processes may need to be developed in order for the child to detect emotional content in a face and what does the emergence of this facial configuration mean. In this sense, although visual acuity plays an important role in perceptual discrimination, it is important to control whether facial configurations that represent the same "emotional" category or affective attribute are processed as such, which also involves the maturation of subcortical and cortical structures that take part in emotion-related processing (Acerra, Burnod, & de Schonen, 2002, Braun, 2011; Joseph, 1999).

Some studies have addressed the question of the emotional meaning of stimuli by relying on adaptation protocols that used different models with the same facial configurations (e.g., Nelson, Morse, & Leavitt, 1979), or different intensities of the same "emotion" during the adaptation phase (e.g., Ludemann & Nelson, 1988). The use of adaptation paradigms allows examining whether the target stimulus is considered the same as the adaptor in some aspect. In addition, it allows one to determine whether there are order effects dependent on the adaptation. In these studies, 6- and 7-month-old infants consistently discriminated between *happy* and other facial configurations when they were adapted to *happy* faces but not when they were adapted to other facial configurations (Caron, Caron, & Myers, 1982; Kotsoni, de Haan, & Johnson, 2001; Ludemann & Nelson, 1988; Nelson & Dolgin, 1985). Studies examining neural responses using event-related potentials with 7-month-old infants also show differential neural activation to *happy* faces vs. other "emotional" categories (e.g., Grossmann, Striano, & Friederici, 2007; Nelson & de Haan, 1996). The results between other categories, namely between *fear* and *anger*, are not consistent (e.g., de Haan, Belsky, Reid, Volein & Johnson, 2004; Kobiella, Grossmann, Reid, & Striano, 2006; Nelson & de Haan, 1996). This has been explained by the fact that these configurations refer to "emotions" with the same valence and less experienced in a normal environment (Malatesta & Haviland, 1982).

The facilitated processing of *happy* faces has been called “*happiness* advantage” and may be due to the higher prevalence of *happy* facial configurations experienced by the infants at these ages. This effect is also consistent with PP: for low-risk samples of infants, facial configurations of *happiness* are likely to be frequently present in the visual field of the child. Since perceptual priors are dependent on previous experiences of the world, the fact that configurations of *happiness* are regularly expected and confirmed makes this prior a more robust belief. Therefore, in adaptation paradigms, the habituation to a facial configuration that is already a prior, increases not only the P(B), but also the P(E|B).

Conversely, samples of infants at-risk present differential behavioral and neural patterns to facial “expressions” of emotion, stressing the effect of the child’s experiences on the priors (for a systematic review of the effects of early neglect on emotion recognition, see Doretto & Scivoletto, 2018). Curtis and Cicchetti (2013) studied brain activity to facial configurations in 15-month-old maltreated and non-maltreated infants. The authors found increased neural responses to "affective" novel stimuli in both samples, which is expected according to the PP model, but the effect of novelty differed between samples: the maltreated infants had higher amplitudes for *happy* faces whereas the non-maltreated group was more reactive to *angry* faces. These findings may be explained by the existence of different priors for facial configurations in these different groups of children.

Other effects of differential exposure to facial configurations are reported in the developmental literature. Pre-schoolers subjected either to neglect or to abuse show differential patterns in their abilities to recognize and discriminate facial configurations: neglected children seem to be less accurate in discriminating a variety of facial configurations, whereas abused children show a specific bias to *angry* facial cues (Pollak, Cicchetti, Hornung, & Reed, 2000). In line with these findings, physically abused school-age children required less perceptual information to recognize cues of *anger* in facial configurations than control children, which the authors interpreted as a facilitated access to a mental representation of the "expression of" *anger* (Pollak, Messner, Kistler, & Cohn, 2009; Pollak & Sinha, 2002). This bias to *angry* facial configurations is also corroborated by data on their physiological reactivity (Pollak, Cicchett, Klorman, & Brumaghim, 1997; Pollak, Klorman, Thatcher, & Cicchetti, 2001; Pollak & Tolley-Schell, 2002). Another study revealed that maltreated children, between 8- and 15 years of age-, show faster reaction times than non-maltreated children in the recognition of *fearful* facial configurations with less intense manifestations of the "expression" (Masten et al., 2008). A study with abused children, between 6 and 17 years of age, found significant deficits in the inference of positive emotions in a set of photos of "expressive" human eyes but no difficulties in the inference of negative emotions (Koizumi & Takagishi, 2014).

More studies regarding the processing of emotion-related facial configurations in the different types of abuse and neglect would provide a clearer picture of how these experiences impact in the predictive processing of these facial configurations, especially regarding the reduced discrimination ability in neglected children. In addition, it is not always clear whether the results of the studies reviewed depend solely on discrimination and labeling of facial configurations. It is likely that emotional interpretation of the facial stimuli accounts for some of the findings, especially in samples of older children and adolescents.

Comparing the results of low-risk and at-risk samples suggests that different experiences during development lead to different abilities to process different facial configurations and interpret their emotional meaning. We may speculate that these differences occur because of differential exposure to facial expression stimuli and the contexts in which they are viewed – or in PP terms to the establishment of different priors. Studies exploring the natural statistics of visual experience in infants (e.g., Jayaraman et al., 2017; Smith et al., 2018) will likely prove fruitful in examining these effects. Furthermore, even in low-risk samples, despite the robustness of the "*happiness* advantage" in the literature, it is still unclear whether this is an emotional or merely a perceptual effect. If it is perceptual, it may be the case that the average face of the "face space" has configurative properties, that resemble a *happy* configuration without an actual awareness of the emotional content and meaning. However, if the child is able to process the emotional and affective content of the facial configuration, then configurations with similar affective properties such as pleasant valence, positive outcomes, and similar contextual cues may evoke the same pattern. Some inconsistent findings in children represent this effect, especially regarding the inability to distinguish between certain emotional categories unless they are portrayed by the primary caregiver or a close relative (e.g., Kahana-Kalman & Walker-Andrews, 2001). Indeed, there may not be a clear understanding of the emotional content, but there is already an association between that facial configuration a specific person and possible outcomes for the child. In addition, it is likely that there is some generalization to other perceptually similar agents, but not a complete emotional interpretation of the facial configuration. However, the existing literature regarding emotion-related processing based on facial configurations is dominated by the comparison of different “emotional” categories or intensities driven by the assumption that facial “expressions” are simple consequences of internal emotional states and therefore indicators of the person’s emotions. This approach has not fully appreciated the distinction between the development of abilities to process facial configurations and the ability to infer and understand the emotional states of others. Next, we examine this theoretical position and critically contrast it with alternative models.

From a valence-based processing to the establishment of emotional categories

Research on emotion-related processing is highly dependent on the concept of emotion itself. As mentioned by Gross and Barrett (2011), all theories agree that emotions include subjective experiences, specific physiological responses, and features of expressive behavior, but their core assumptions firmly differ. Basic Emotion Models (BEM) represent the greater tradition in emotion research and consider each emotion to have a discrete identity with a specific cause, a unique pattern of physiological and facial responses, and particular neural processing mechanisms. Competing perspectives, such as Psychological Construction Models (PCM), consider emotion as a state that is composed of various ingredients and their associated manifestations (Barrett, 2013). Most PCM consider valence/pleasantness and activation/arousal as fundamental aspects of core affect (Kuppens, Tuerlinckx, Russell, & Barrett, 2013; Russell, 2009). Some theories add other dimensions, but valence and activation remain as fundamental ones for the occurrence of affective states, even if these are not the only dimensions involved (Carrol, Yik, Russell, & Barrett, 1999). In terms of neural mechanisms of both emotional experience and emotion-related processing, BEM and PCM differ to the extent that the former argues that there is specific processing of each emotional category, while the latter considers to be a specific processing of each dimension instead (Hamann, 2012; Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012). It is important to stress that these models do not merely discuss the role of these affective properties but rethink the developmental and conceptual aspects of the generation of emotional processes themselves (for an in-depth discussion see Barrett, 2017b).

Currently, studies in the field of facial "expression" of emotion are mostly based on BEM, which influences what is known about emotion-related processing of these stimuli. As we have stressed above, developmental research is not an exception to this phenomenon, but studies point to the occurrence of a combination of dimensional and categorical processing of facial stimuli along development. By 10-months of age, infants can already differentiate groups of "emotional" categories that vary between themselves in valence, that is, between a group of facial configurations typically considered to reflect positive or negative states (Ludemann, 1991). This is consistent with the proposal made by Clark (2013) regarding belief development, in that the affective significance of facial configurations is first processed in a broader way, in this case firstly guided by valence. The processing of valence is an important aspect of behavioral management regarding approach or avoidance strategies in response to desirable (positive, associated with rewards) or harmful stimuli (negative, associated with punishment; see Bach & Dayan, 2017). Indeed, the ability to understand social cues that signal rewarding or punishing outcomes is fundamental to pre-verbal infants.

Widen (2013) reviews the literature on children's interpretation of facial configurations between 2 and 9 years of age and proposes the "broad-to-differentiated hypothesis". She suggests that children slowly start to acquire emotional categories, first by being sensitive to the valence associated with facial stimuli, and then by obtaining the awareness of the different categories according to cultural and social influences. This gradual differentiation is not specific for the emotion-related processing of facial configurations, but for emotional development in general (Bridges, 1932), which is also postulated as a combination of inherent factors and environmental influences, such as routines and stimulation patterns.

Two- and 3-year-old children seem to categorize a range of pleasant and unpleasant emotions from facial configurations as “happy” and “angry”, respectively (Russell & Widen, 2002), evolving to a more discriminated version of negative configurations, specifically to an “angry” labeling for configurations of *anger* and *disgust*, and a “sadness” labeling for *sad* and *fearful* faces (Widen, 2013). This is illustrative of the maintenance of a form of valence-based discrimination, in which children have more difficulty in discriminating between emotional faces with the same valence and similar arousal levels until stabilizing in what adults consider to be the correct labels of basic emotions by 80 months (Widen, 2013, 2016). In adolescence and adulthood, the focus on a categorical approach is still very prominent (see Batty & Taylor, 2006; Taylor, Batty, & Itier, 2004), essentially because there is a stabilization of recognition abilities, which somewhat corroborates the categorical perspective. The same is true for adult studies of facial “expressions” of emotion, despite evidence suggesting that neural activity in the visual cortex is correlated with the affective arousal of facial configurations regardless of “emotional” category (Almeida et al., 2016).

In summary, literature suggests that the processing of facial “expressions” evolves from perceptual discrimination of facial configurations to a valence-based differentiation until the establishment of “emotional” categories by adolescence. These results match Clark’s (2013) description of the generation of representation from perception to a progressive organization into global patterns, becoming more accurate and adding more details throughout development. This change from broad, valence-based processing of facial configurations goes hand-in-hand with the development of other phenomena; for example, emotional regulation

(Martin & Ochsner, 2016), the sensitivity to understand and label one's own emotional states (Dennis, Malone, & Chen, 2009), social processes such as theory of mind (Widen, 2016), and the acquisition of contextual cueing (Baldwin & Moses, 1996).The preponderance of valence in emotion-related processing during the first years suggests the need to examine how infants and children extract emotional information from facial configurations beyond the discrimination or recognition of "emotional" categories. In this sense, studies on the behavioral and physiological responses of infants and toddlers to facial configurations that vary in their dimensional affective properties (i.e., valence and arousal), and not in "emotional" category *per se*, may provide a detailed timeline of the development of their recognition ability.

A predictive perspective on the processing of emotion-related facial configurations

In the present work, we have described the general aspects of PP and have explored existing research on the development of facial processing and emotion-related processing of facial configurations. Next, we will attempt to integrate the research findings reviewed above within a PP framework.

Early preferences of fetuses and newborns for face-like stimuli are consistent with a very early and unspecific prior that may facilitate visual exposure to faces (Johnson, 2011). With this visual experience, this prior may evolve from a primitive form to an average face that is highly influenced by the facial configurations that are more frequent in the child's daily life. With the accrual of visual experience with facial configurations, a "face space" may develop, which is an array composed of the main variations that the average face can assume. The available evidence suggests that this average face begins by resembling the features of a *happy* facial configuration in children with typical development explaining the behavioral preferences and patterns of neural activity that define the "happiness advantage" effect. At this point, the processing of facial “expressions” is likely dominated by perceptual discrimination without a full understanding of emotional meanings. We hypothesize that, beyond the physical discrimination of facial configurations, children may begin to broadly understand the valence of these configurations. This is possibly guided by the value of the outcomes of these expressions for children (e.g., a facial expression of *anger* or *sadness* by their caregivers may mean that they will be punished and thus become associated with negative outcomes). However, children are not able to discriminate each “emotional” category individually, also observed when labeling facial configurations of similar negative valence as “anger”, for example. Progressively, different “subface spaces” may be developed for each main emotional category. Each subface space would have its own prototype (i.e., average configuration per category). This differentiation may occur due to the creation of more precise associations between: (a) the specific facial configurations, (b) the contexts in which they arise (e.g., although *happy* expressions are the most seen in a low-risk environment, a *happy* face may not be expected in a context of grief or after the child has done something wrong), and (c) their outcomes for the person exhibiting the facial configuration (e.g., unpleasant behaviors) and for the perceiver (e.g., punishments). The ability to create more precise associations between facial configurations, contexts, and outcomes is facilitated by cognitive development, such as language improvement, mentalizing, and inferential abilities, allowing a better understanding of emotional states that may accompany facial displays.

Once the child has established a distinction between “emotional” categories of facial displays, we can consider that viewing a facial expression may elicit low-level perceptual PEs, expectably proportional to the amount of deviation from the perceptual prior, and high-level PEs, related to the adequacy of the emotion to the wider context. Perceptually, different levels of activation/arousal (or extreme levels of valence) may be a factor that modulates the lower-level PEs after adolescence and in adulthood. More aroused faces reflect more muscle activity and, therefore, these stimuli may deviate from the perceptual average of that subface space. In addition, facial expressions may also elicit higher-order PEs if the specific valence or “emotion displayed” does not match the high-level predictions regarding the present context. This may be implemented in a predictive hierarchy as high-level priors that constrain the lower-level predictions: for instance, safe and familiar environments may induce a perceptual prior that is biased towards pleasant facial configurations whereas threatening environments may lead to predictions that pleasant facial configurations would not be present. In this example, the same pleasant configuration (i.e., the same physical stimulus) would induce a PE if viewed in a threatening environment, but not in a safe environment.

Finally, the assumption that processing emotion-related facial configuration follow the tenets of PP implies that our conceptualization of emotional development should include (a) highlighting the statistical and quantitative influence of the environment in the development of beliefs, (b) stressing the influence of early social experiences, and (c) emphasizing the multiple levels at which internal and external factors contribute to development. In this sense, we describe a convergence between these predictive processes and the ontogenetic course of face processing and emotion-related processing of facial configurations, suggesting that hypotheses regarding developmental mechanisms may be proposed based on recent rethinking the core properties of brain function at each stage.

Future Directions

In the present work, no computational models were actually tested or discussed in-depth. The PP framework includes a computational understanding of how the generation of predictions may occur and how internal representations may be updated, which may raise novel hypotheses. However, our goal was to argue for the complementary nature of the processing of emotion-related facial configurations and these predictive processes and to stimulate future efforts in considering PP in development. In this sense, this review uncovers the need for computational models that are able to disclose the ways in which predictive processes appear and change in the developing brain in a wide range of low- and high-level processes. Specifically, adopting a PP view raises potentially novel questions in development: how are priors initially represented and developed?; are the mechanisms for perceptual learning similar across development?; or do these mechanisms change, potentially by implementing different algorithms at different developmental stages?

Furthermore, while the emotion-related processing of facial configurations may be approached as a single specific phenomenon, we know that it is a complex, multi-level and multimodal process, dependent on the context. This complexity should be considered methodologically, through a more careful understanding of the impact of stimulus properties on the outcomes and the conclusions derived. Specifically, in the present work, we reviewed how the affective properties of the face (valence and arousal) may redefine the way we believe our brain to process facial "expressions" of emotion throughout development. However, other properties should be taken into account in the search for more ecological validity in emotion-related processing studies, such as the use of dynamic facial stimuli and the consideration of contextual and multimodal cues (Aviezer et al., 2008; Aviezer, Dudarev, Bentin, & Hassin, 2011; Barrett, Mesquita & Gendron, 2011; Grossmann, 2010).

The complexity of emotion processing also needs to be considered conceptually, in order to continue with the revision of concepts such as facial "expressions" of emotion, a reform that has become increasingly active over the last few years (Barrett, 2017b; Gendron, Crivelli, & Barrett, 2018). Moreover, in the present review we addressed the evidence regarding the processing of facial expressions mainly as an exteroceptive phenomenon, that is, as a product of visual perception and emotional understanding. However, it is known that interoception plays a role in perceptual and emotional processes (Allen & Friston, 2016; Atzil

& Barrett, 2017; Barrett, 2017b). We have only addressed this issue superficially, but it may open a promising avenue for research in this field.

As a conclusion, regarding the emotion-related processing of facial configurations as a compromise between expectations (priors) and sensory inputs that dynamically shape those expectations (via integration of PEs) from early stages of development leads to rethinking the developing brain, reframes the special perceptual significance of the face, and empowers top-down mechanisms in understanding emotional development in a growing social agent.

References

Acerra, F., Burnod, Y., & de Schonen, S. (2002). Modelling aspects of face processing in early infancy. *Developmental Science, 5*, 98-117. https://doi.org/10.1111/1467-7687.00215

Adams, R. A., Brown, H. R., & Friston, K. J. (2014). Bayesian inference, predictive coding and delusions. *Avant, 5,* 51-88. https://doi.org/10.12849/50302014.0112.0004

Ainsworth, M. D., & Bell, S. M. (1970). Attachment, exploration, and separation: Illustrated by the behavior of one-year-olds in a strange situation. *Child Development*, *41*(1), 49-67. http://doi.org/ 10.2307/1127388

Aitchison, L., & Lengyel, M. (2017). With or without you: Predictive coding and Bayesian inference in the brain. *Current Opinion in Neurobiology, 46*, 219-227. https://doi.org/10.1016/j.conb.2017.08.010

Allen, M., & Friston, K. J. (2016). From cognitivism to autopoiesis: Towards a computational framework for the embodied mind. *Synthese, 195*, 2459-2482. https://doi.org/10.1007/s11229-016-1288-5

Almeida, P. R., Ferreira-Santos, F., Chaves, P. L., Paiva, T. O., Barbosa, F., & Marques-Teixeira, J. (2016). Perceived arousal of facial expressions of emotion modulates the N170, regardless of emotional category: Time domain and time-frequency dynamics. *International Journal of Psychophysiology, 99*, 48-56. https://doi.org/10.1016/j.ijpsycho.2015.11.017

Apps, M. A., & Tsakiris, M. (2013). Predictive codes of familiarity and context during the perceptual learning of facial identities. *Nature Communications, 4*, 2698. https://doi.org/10.1038/ncomms3698

Atzil, S., & Barrett, L. F. (2017) Social regulation of allostasis: Commentary on "Mentalizing homeostasis: The social origins of interoceptive inference" by Fotopoulou and Tsakiris. *Neuropsychoanalysis, 19*, 29-33. https://doi.org/10.1080/15294145.2017.1295214

Aviezer, H., Dudarev, V., Bentin, S., & Hassin, R. R. (2011). The automaticity of emotional face-context integration. *Emotion, 11*(6)., 1406-1414. https://doi.org/10.1037/a0023578.

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M., & Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science, 19*(7), 724-732. https://doi.org/10.1111/j.1467-9280.2008.02148.x

Bach, D. R., & Dayan, P. (2017). Algorithms for survival: A comparative perspective on emotions. *Nature Reviews Neuroscience, 18*, 311-319. https://doi.org/10.1038/nrn.2017.35

Balas, B. (2012). Bayesian face recognition and perceptual narrowing in face-space. *Developmental Science*, *15*(4), 579–588. http://doi.org/10.1111/j.1467-7687.2012.01154.x

Baldwin, D. A., & Moses, L. J. (1996). The ontogeny of social information gathering. *Child Development, 67*, 1915-1939. https://doi.org/10.2307/1131601

Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and nurture in own-race face processing. *Psychological Science, 17*, 159-163.

Barrera, M. E., & Maurer, D. (1981). The perception of facial expressions by three-month-olds. *Child Development, 52*, 203-206. https://doi.org/10.2307/1129231

Barrett, L. F. (2013). Psychological construction: The Darwinian approach to the science of emotion. *Emotion Review, 5*, 379-389. https://doi.org/10.1177/1754073913489753

Barrett, L. F. (2017a). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience, 12*, 1-23. https://doi.org/10.1093/scan/nsw154

Barrett, L. F. (2017b). *How emotions are made: The secret life of the brain*. Boston: Houghton Mifflin Harcourt.

Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. Current *Directions in Psychological Science, 20*(5), 286-290. https://doi.org/10.1177/0963721411422522

Barrett, L. F., & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience, 16*(7), 419-429. https://doi.org/10.1038/nrn3950

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

Batty, M., & Taylor, M. J. (2006). The development of emotional face processing during childhood. *Developmental Science, 9*, 207-220. https://doi.org/10.1111/j.1467-7687.2006.00480.x

Bolstad, W. M., & Curran, J. M. (2017). *Introduction to Bayesian statistics (3rd ed.)*. Hoboken, NJ: John Wiley & Sons.

Braddick, O. (2017). Human development: Faces in the womb. *Current Biology, 27*(14), PR704-R706. https://doi.org/10.1016/j.cub.2017.06.014

Braun, K. (2011). The prefrontal-limbic system: Development, neuroanatomy, function, and implications for socioemotional development. *Clinics in Perinatology, 38*, 685-702, https://doi.org/ 10.1016/j.clp.2011.08.013

Bridges, K. M. B. (1932). Emotional development in early infancy. *Child Development, 3,* 324-341.

Brodski, A., Paasch G.-F., Helbling, S., & Wibral, M. (2015). The faces of predictive coding. *Journal of Neuroscience, 35*, 8997-9006. https://doi.org/10.1523/JNEUROSCI.1529-14.2015

Brodski-Guerniero, A., Paasch, G.-F., Wollstadt, P., Özdemir, O., Lizier, J. T., & Wibral, M. (2017). Information-theoretic evidence for predictive coding in the face-processing system. *Journal of Neuroscience, 37*, 8273-8283. https://doi.org/10.1523/JNEUROSCI.0614-17.2017

Bruce, V., & Young, A. (1986). Understanding race recognition. *British Journal of Psychology, 77*, 305-327. https://doi.org/10.1111/j.2044-8295.1986.tb02199.x

Caron, R. F., Caron, A. J., & Myers, R. S. (1982). Abstraction of invariant face expressions in infancy. *Child Development, 53*, 1008-1015. https://doi.org/10.2307/1129141

Carrol, J. M., Yik, M. S. M., Russell, J. A., & Barrett, L. F. (1999). On the psychometric principles of affect. *Review of General Psychology, 3*, 14-22. https://doi.org/10.1037/1089-2680.3.1.14

Cassia, V. M., Turati, C., & Simion, F. (2004). Can a non-specific bias toward top-heavy patterns explain newborns' face preference? *Psychological Science, 15*(6), 379-383. https://doi.org/10.1111/j.0956-7976.2004.00688.x

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*, 181-204. https://doi.org/10.1017/s0140525x12000477

Clark, A. (2015). Embodied Prediction. In T. Metzinger & J. M. Windt (Eds). *Open MIND*: 7(T). Frankfurt: MIND Group. https://doi.org/10.15502/9783958570115

Cohen-Kadosh, K., & Johnson, M. H. (2007) Developing a cortex specialized for face perception. *Trends in Cognitive Science, 11*, 267-269. https://doi.org/10.1016/j.tics.2007.05.003

Curtis, W. J., & Cicchetti, D. (2013). Affective facial expression processing in 15-month-old infants who have experienced maltreatment: An eventrelated potential study. *Child Maltreatment, 18,* 140–154. https://doi.org/10.1177/1077559513487944.

de Haan, M., Belsky, J., Reid, V., Volein, A., & Johnson, M. H. (2004). Maternal personality and infants' neural and visual responsivity to facial expressions of emotion. *Journal of Child Psychology and Psychiatry, 45*, 1209-1218. https://doi.org/10.1111/j.1469-7610.2004.00320.x

de Haan, M., Johnson, M. H., & Halit, H. (2003). Development of face-sensitive event-related potentials during infancy: A review. *International Journal of Psychophysiology, 51*, 45-58. https://doi.org/10.1016/S0167-8760(03)00152-1

de Haan, M., Pascalis, O., & Johnson, M. H. (2002). Specialization of neural mechanisms underlying face recognition in human infants. *Journal of Cognitive Neuroscience*, *14,* 199-209. https://doi.org/10.1162/089892902317236849 de Haan, M., & Nelson, C.A. (1997). Recognition of the mother's face by 6-month-old infants: A neurobehavioral study. *Child Development, 68,* 187-210. https://doi.org/10.1111/j.1467-8624.1997.tb01935.x

de Haan, M., Johnson, M. H., & Halit, H. (2007). Development of face-sensitive event-related potentials during infancy. In M. de Haan (Ed.), *Infant EEG and event-related potentials* (pp. 77-100). New York: Psychology Press.

de Haan, M., Johnson, M. H., Maurer, D., & Perrett, D. I. (2001). Recognition of individual faces and average face prototypes by 1-and 3-month-old infants. *Cognitive Development, 16,* 659-678. https://doi.org/10.1016/S0885-2014(01)00051-X

de Haan, M., & Matheson, A. (2009). The development and neural bases of processing emotion in faces and voices. In M., de Haan, M. Gunnar, (Eds.), *Handbook of Developmental Social Neuroscience* (pp. 107-120). New York: Guilford Press.

de Haan, M., & Nelson, C. A. (1998). Discrimination and categorization of facial expressions of emotion during infancy. In A. Slater (Ed.) *Perceptual development: visual, auditory and language development in infancy*. (pp. 287-309). London, UK: University College London Press.

Denève, S., Alemi, A., & Bourdoukan, R. (2017). The brain as an efficient and robust adaptive learner. *Neuron, 94*(5), 969-977. https://doi.org/10.1016/j.neuron.2017.05.016.

Dennis, T.A., Malone, M. M., & Chen, C. C. (2009). Emotional face processing and emotion regulation in children: An ERP study. *Developmental Neuropsychology, 34,* 85-102. https://doi.org/10.1080/87565640802564887

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology General, 115,* 107-117. https://doi.org/10.1037/0096-3445.115.2.107

Dienes, Z. (2008). *Understanding Psychology as a science: An introduction to scientific and statistical inference*. New York: Palgrave MacMillan.

Doretto, V., & Scivoletto, S. (2018). Effects of early neglect experience on recognition and processing of facial expressions: A systematic review. *Brain Sciences*, *8*(1), 10. http://doi.org/10.3390/brainsci8010010

Everitt, B. S., & Skrondal, A. (2010). *The Cambridge dictionary of Statistics* (4th ed.). New York: Cambridge University Press.

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic? *Emotion Review, 3,* 364–370.

Farroni, T., Chiarelli, A. M., Lloyd-Fox, S., Massaccesi, S., Merla, A., Di Gangi, V., Mattarello, T., Faraguna, D., & Johnson, M. H. (2013). Infant cortex responds to other humans from shortly after birth. *Scientific Reports, 3*, 1–5. https://doi.org/10.1038/srep02851

Farroni T., Johnson M. H., Menon E., Zulian L., Faraguna D., & Csibra G. (2005). Newborns' preference for face-relevant stimuli: effects of contrast polarity. *Proceedings of the National Academy of Sciences, 102*, 17245-17250. https://doi.org/10.1073/pnas.0502205102

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences, 99*, 9602-9605. https://doi.org/10.1073/pnas.152159999

Feinman, S. (1992). *Social referencing and the social construction of reality in infancy*. New York, NY, US: Plenum Press. http:// doi.org/10.1007/978-1-4899-2462-9 Ferreira-Santos, F. (2016). The role of arousal in predictive coding. *Behavioral and Brain Sciences, 39*, e207. https://doi.org/10.1017/S0140525X15001788

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for “top-down” effects. *Behavioral and Brain Sciences*, *39*, e229. https://doi.org/10.1017/S0140525X15000965

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

Friston, K. (2008). Hierarchical models in the brain. *PLoS Computational Biology*, *4*(11), e1000211. https://doi.org/10.1371/journal.pcbi.100021

Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics*, *104*, 137-160. https://doi.org/10.1007/s00422-011-0424-z

Friston, K. J. (2010) The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*, 127-138. https://doi.org/10.1038/nrn2787

Friston, K. J. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, *13*, 293-301. https://doi.org/10.1016/j.tics.2009.04.005.

Friston, K. J. (2012) The history of the future of the Bayesian brain. *Neuroimage*, *62*, 1230-1233. https://doi.org/ 10.1016/j.neuroimage.2011.10.004.

Gaither, S. E., Pauker, K., & Johnson, S. P. (2012). Biracial and monoracial infant own-race face perception: An eye tracking study. *Developmental Science*, *15*, 775-782. https://doi.org/10.1111/j.1467-7687.2012.01170.x

Gallagher, S., & Allen, M. (2016). Active inference, enactivism and the hermeneutics of social cognition. *Synthese*, *195*, 2627-2648. https://doi.org/10.1007/s11229-016-1269-8

Gendron, M., Crivelli, C., & Barrett, L. F. (2018). Universality reconsidered: Diversity in making meaning of facial expressions. *Current Directions in Psychological Science*, *27*, 211–219. https://doi.org/10.1177/0963721417746794

Gładziejewski, P. (2016). Predictive coding and representationalism. *Synthese, 193*, 559-582. https://doi.org/10.1007/s11229-015-0762-9

Goren, C. C., Sarty, M., & Wu, P. Y. K. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics, 56,* 544-549.

Gross, J.J. (2013). Emotion regulation: Taking stock and moving forward. *Emotion, 13,* 359–65. http://doi.org/ 10.1037/a0032135

Gross J. J., & Barrett, L. F. (2011). Emotion generation and emotion regulation: one or two depends on your point of view. *Emotion Review, 3,* 8-16.https://doi.org/10.1177/1754073910380974

Gross, J. J., & Thompson, R. A. (2007). Emotion Regulation Conceptual Foundations. In J. J. Gross (Ed.), *Handbook of Emotion Regulation* (pp. 3-24). New York: Guilford Press.

Grossmann, T. (2010). The development of emotion perception in face and voice during infancy. *Restorative Neurology and Neuroscience, 28,* 219-236. https://doi.org/10.3233/RNN-2010-0499.

Grossmann, T., Striano, T., & Friederici. A. D. (2007). Developmental changes in infants' processing of happy and angry facial expressions: A neurobehavioral study. *Brain and Cognition, 64*, 30-41.

Halit H., de Haan M., Johnson M. H. (2003). Cortical specialization for face processing: face-sensitive event-related potential components in 3- and 12-month-old infants. *Neuroimage, 19*, 1180-1193. https://doi.org/10.1016/S1053-8119(03)00076-4

Halit, H., Csibra, G., Volein, A., & Johnson, M.H. (2004). Face-sensitive cortical processing in early infancy. *Journal of Child Psychology and Psychiatry, 45*, 1228-1234.

Hamann, S. (2012). Mapping discrete and dimensional emotions onto the brain: controversies and consensus. *Trends in Cognitive Sciences, 16,* 458-466. https://doi.org/10.1016/j.tics.2012.07.006

Hayden, A., Bhatt, R. S., Reed, A., Corbly, C. R., & Joseph, J. E. (2007). The development of expert face processing: Are infants sensitive to normal differences in second-order relational information? *Journal of Experimental Child Psychology, 97*, 85-98. https://doi.org/10.1016/j.jecp.2007.01.004

Hohwy, J. (2013). *The predictive mind*. Oxford: Oxford University Press.

Huang, Y., & Rao, R. P. N. (2011). Predictive coding. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*, 580-593. https://doi.org/ 10.1002/wcs.142.

Izard, C.E. (1977). *Human Emotions*. New York: Plenum.

Jayaraman, S., Fausey, C. M., & Smith, L. B. (2017). Why are faces denser in the visual experiences of younger than older infants? *Developmental Psychology, 53*, 38-49. https://doi.org/10.1037/dev0000230

Johnson, M. H. (2000). Functional brain development in infants: elements of an interactive specialization framework. *Child Development, 71*(1), 75-81. https:// doi.org/10.1111/1467-8624.00120

Johnson, M. H. (2001). Functional brain development in humans. *Nature Reviews in Neuroscience, 2*, 475-483. https://doi.org/10.1038/35081509

Johnson, M. H. (2005). Subcortical face processing. *Nature* Reviews *Neuroscience, 6*, 766-773. https://doi.org/10.1038/nrn1766

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40*, 1-19. https://doi.org/10.1016/0010-0277(91)90045-6

Johnson, M. H. (2011). Face perception: A developmental perspective. in A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *The Oxford handbook of face perception* (pp. 3-14). Oxford: Oxford University Press.

Joseph, R. (1999). Environmental influences on neural plasticity, the limbic system, emotional development and attachment: A review. *Child Psychiatry and Human Development, 29*, 189-208. https://doi.org/10.1023/A:1022660923605

Kahana-Kalman, R., & Walker-Andrews, A. S. (2001). The role of person familiarity in young infants' perception of emotional expression. *Child Development, 72*, 352-369. https://doi.org/10.1111/1467-8624.00283

Kaneshige, T., & Haryu, E. (2015) Categorization and understanding of facial expressions in 4-month-old infants. *Japanese Psychological Research, 57*, 135-142.

Kelly, D. J., Liu, S., Ge, L., Quinn, P. C., Slater, A. M., Lee, K., Liu, Q, & Pascalis, O. (2007). Cross-race preferences for same-race faces extend beyond the African versus Caucasian contrast in 3-month-old Infants. *Infancy, 11*, 87-95. https://doi.org/10.1080/15250000709336871

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neuroscience, 27,* 712-719. https://doi.org/10.1016/j.tins.2004.10.007

Kobiella, A., Grossmann, T., Reid, V. M., & Striano, T. (2008). The discrimination of angry and fearful facial expressions in 7-month-old infants: An event-related potential study. *Cognition and Emotion, 22,* 134-146. https://doi.org/10.1080/02699930701394256

Koizumi, M., & Takagishi, H. (2014). The Relationship between Child Maltreatment and Emotion Recognition. *PLoS ONE, 9*(1), e86093. http://doi.org/10.1371/journal.pone.0086093

Kok P., & de Lange F. P. (2014) Shape perception simultaneously up- and downregulates neural activity in the primary visual cortex. *Current Biology, 24*, 1531-1535. https://doi.org/ 10.1016/j.cub.2014.05.042.

Kotsoni, E., de Haan, M., & Johnson, M. H. (2001). Categorical perception of facial expressions by 7-month-old infants. *Perception, 30*, 1115-1125. https://doi.org/10.1068/p3155

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological Bulletin, 139*, 917-940. https://doi.org/10.1037/a0030811

Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *The Quarterly Journal of Experimental Psychology, 53*, 513-536. https://doi.org/10.1080/027249800390583;

Lerner, R. M., Hershberg, R. M., Hilliard, L. J., & Johnson, S. K. (2015). Concepts and theories of Human Development. In M. H. Bornstein & M. E. Lamb (Eds.), *Developmental Science: An Advanced Textbook* (7th ed.). New York: Psychology Press.

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences, 35*, 121-143. https://doi.org/10.1017/S0140525X11000446

Lochmann, T., & Deneve, S. (2011). Neural processing as causal inference. *Current Opinion in Neurobiology, 21*(5), 774-781. https://doi.org/10.1016/j.conb.2011.05.018.

Ludemann, P. M., & Nelson, C. A. (1988). The categorical representation of facial expressions by 7-month-old infants. *Developmental Psychology, 24,* 492-501. https://doi.org/10.1037//0012-1649.24.4.492.

Ludemann, P. M. (1991). Generalized discrimination of positive facial expressions by seven- and ten-month-old infants. *Child Development, 62,* 55-67. https://doi.org/10.2307/1130704

Malatesta, C. Z., & Haviland, J. M. (1982). Learning display rules: The socialization of emotion expression in infancy. *Child Development, 53,* 991-1003. https://doi.org/10.2307/1129139

Martin, R. E., & Ochsner, K. N. (2016). The neuroscience of emotion regulation development: implications for education. *Current Opinion in Behavioral Sciences, 10,* 142-148. https://doi.org/10.1016/j.cobeha.2016.06.006

Masten, C. L., Guyer, A. E., Hodgdon, H. B., McClure, E. B., Charney, D. S., Ernst, M., … Monk, C. S. (2008). Recognition of facial emotions among maltreated children with high rates of post-traumatic stress disorder. *Child Abuse & Neglect*, *32*(1), 139–153. http://doi.org/10.1016/j.chiabu.2007.09.006

Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biological Cybernetics, 66,* 241-251.

Nelson C.A. (2001). The development and neural bases of face recognition. *Infant and Child Development. 10*, 3-18. https://doi.org/10.1002/icd.239

Nelson C. A. (2003). The developmental of face recognition reflects an experience-expectant and experience-dependent process, in O. Pascalis & A. Slater (Eds.), *The development of face processing in infancy and early childhood: Current perspectives* (pp. 79-88) New York: Nova Science Publishers.

Nelson, C. A., & Dolgin, K. (1985). The generalized discrimination of facial expressions by 7-month-old infants. *Child Development*, *56*, 58-61. https://doi.org/10.2307/1130173

Nelson, C. A., Morse, P. A., & Leavitt, L. A. (1979). Recognition of facial expressions by 7-month-old infants. *Child Development, 50,* 1239-1242. https://doi.org/10.2307/1129358

Nelson, C.A., & de Haan, M. (1996). Neural correlates of infants' visual responsiveness to facial expressions of emotion. *Developmental Psychobiology, 29,* 577-595. https://doi.org/10.1002/(SICI)1098-2302(199611)

Perfors, A., Tenenbaum, J. B., Griffiths, T. L., & Xu, F. (2011). A tutorial introduction to Bayesian models of cognitive development. *Cognition*, *120*(3), 302–321. https://doi.org/10.1016/j.cognition.2010.11.015

Pollak, S. D., Cicchetti, D., Hornung, K., & Reed, A. *(*2000) Recognizing emotion in faces: Developmental effects of child abuse and neglect*.* *Developmental Psychology, 36,* 679-688. https://doi.org/10.1037//0012-1649.36.5.679.

Pollak, S. D., Klorman, R., Thatcher, J. E., & Cicchetti, D. (2001). P3b reflects maltreated children's reactions to facial displays of emotion. *Psychophysiology, 38*(2), 267-274. https:// doi.org/10.1017/S0048577201990808

Pollak, S. D., Messner, M., Kistler, D. J., & Cohn, J. F. (2009). Development of perceptual expertise in emotion recognition. *Cognition*, *110*(2), 242-247. https://doi.org/10.1016/j.cognition.2008.10.010

Pollak, S. D., & Sinha, P. (2002). Effects of early experience on children's recognition of facial displays of emotion. *Developmental Psychology*, *38*(5), 784-791. http://doi.org/10.1037/0012-1649.38.5.784

Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception, 31,* 1109-1121 https://doi.org/10.1068/p3331

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

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

Reid, V. M., Dunn, K. J., Young, R. J., Amu, J., Donovan, T., & Reissland, N. (2017). The human fetus preferentially engages with face-like visual stimuli. *Current* Biology, 27, 1825-1828. https://doi.org/10.1016/j.cub.2017.06.036

Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cognition and Emotion, 23*, 1259-1283. https://doi.org/10.1080/02699930902809375

Russell, J. A., & Widen, S. C. *(*2002). Words versus faces in evoking children's knowledge of the causes of emotions. *International Journal of Behavioral Development*, *26*, 97-103. https://doi.org/10.1080/ 01650250042000582

Schoot, R. van de, Kaplan, D., Denissen, J., Asendorpf, J. B., Neyer, F. J., & Aken, M. A. G. van. (2014). A gentle introduction to Bayesian analysis: Applications to developmental research. *Child Development, 85*(3), 842–860. https://doi.org/10.1111/cdev.12169

Shadmehr, R., Smith, M. A., & Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annual Review Neuroscience, 33*, 89-108. https://doi,org/ 10.1146/annurev-neuro-060909-153135.

Simion, F., & Giorgio, E. D. (2015). Face perception and processing in early infancy: inborn predispositions and developmental changes. *Frontiers in psychology, 6*, 969. http://doi.org/10.3389/fpsyg.2015.00969

Simion, F., Leo, I., Turati, C., Valenza, E., & Dalla Barba, B. (2007) How face specialization emerges in the first months of life. *Progress in Brain Research*, *164*,169-185. https://doi.org/10.1016/S0079-6123(07)64009-6.

Slater, A., Quinn, P.C., Kelly, D.J., Lee, K., Longmore, C.A., McDonald P.R., & Pascalis, O. (2010). The shaping of the face space in early infancy: becoming a native face processor. *Child Development Perspectives, 4*(3), 205-211. http:// doi.org/10.1111/j.1750-8606.2010.00147.x

Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The developing infant creates a curriculum for statistical learning. *Trends in Cognitive Sciences, 22*(4), 325-336. https://doi.org/10.1016/j.tics.2018.02.004

Spratling, M. W. (2017). A review of predictive coding algorithms. *Brain and Cognition, 112*, 92-97. https://doi.org/ 10.1016/j.bandc.2015.11.003.

Spruit, L. (2008). Renaissance views of active perception. In S. Knuuttila and P. Kärkkäinen (Eds), *Theories of perception in medieval and early modern philosophy* (pp. 203-224). New York: Springer.

Sroufe, L. A. (1996) *Emotional Development: The organization of emotional life in the early years*. Cambridge University Press, New York.

Sterling, P. (2012). Allostasis: A model of predictive regulation. *Physiology & Behavior, 106*, 5-15. https://doi.org/10.1016/j.physbeh.2011.06.004

Taylor, M. J., Batty, M., & Itier, R. J. (2004). The faces of development: A review of early face processing over childhood. *Journal of Cognitive Neuroscience, 16,* 1426-1442. https://doi.org/10.1162/0898929042304732

Trainor, L. J. (2012). Predictive information processing is a fundamental learning mechanism present in early development: Evidence from infants. *International Journal of Psychophysiology, 83,* 256-258. https://doi.org/10.1016/j.ijpsycho.2011.12.008.

Trapp, S., Schweinberger, S. R., Hayward, W. G., & Kovács, G. (2018). Integrating predictive frameworks and cognitive models of face perception. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-018-1433-x

Tzourio-Mazoyer, N., de Schonen, S., Crivello, F., Reutter, B., Aujard, Y. & Mazoyer, B. (2002). Neural correlates of woman face processing by 2-month-old infants. *Neuroimage, 15*, 454-461.

van de Schoot, R., Winter, S. D., Ryan, O., Zondervan-Zwijnenburg, M., & Depaoli, S. (2017). A systematic review of Bayesian articles in psychology: The last 25 years. *Psychological Methods*, *22*(2), 217–239. https://doi.org/10.1037/met0000100

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology, 43,* 161-204. https://doi.org/10.1080/14640749108400966

Van den Boomen, C., Munsters, N. M., & Kemner, C. (2017). Emotion processing in the infant brain: The importance of local information. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2017.09.006.

Varela, F. J., Thompson, E., & Rosch, E. (1993). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.

Vincent, B. T. (2015). A tutorial on Bayesian models of perception. *Journal of Mathematical Psychology, 66,* 103–114. https://doi.org/10.1016/j.jmp.2015.02.001

Widen, S. C. (2016). The development of children's concepts of emotion. In L. F. Barrett, M. Lewis, J. M., & Haviland-Jones (Eds.), *Handbook of emotions* (4th ed., pp. 307-318). New York: Guilford Press.

Widen, S. C. (2013). Children's interpretation of facial expressions: The long path from valence-based to specific discrete categories. *Emotion Review, 5,* 72-77. https://doi.org/10.1177/1754073912451492

Young-Browne, G., Rosenfeld, H. M., & Horowitz, F. D. (1978). Infant discrimination of facial expression. *Child Development, 48*, 555-562. https://doi.org/10.2307/1128653