Moritz Köster | Ezgi Kayhan | Miriam Langeloh | Stefanie Hoehl

Making Sense of the World: Infant Learning From a Predictive Processing Perspective

Suggested citation referring to the original publication:
Perspectives on Psychological Science 15 (2020) 3, pp. 562 - 571
DOI: https://doi.org/10.1177/1745691619895071
ISSN: 1745-6916, 1745-6924

Journal article | Version of record

Perspectives on Psychological Science
2020, Vol. 15(3) 562–571

DOI: 10.1177/1745691619895071
www.psychologicalscience.org/PPS

Making Sense of the World: Infant Learning From a Predictive Processing Perspective

Moritz Köster[1,2,3], Ezgi Kayhan[1,4], Miriam Langeloh[1,5], and Stefanie Hoehl[6]

[1]Max Planck Institute for Human Cognitive and Brain Sciences; [2]Faculty of Education and Psychology, Freie Universität Berlin; [3]Department of Psychology, Graduate School of Letters, Kyoto University; [4]Department of Psychology, University of Potsdam; [5]Department of Psychology, Heidelberg University; and [6]Department of Developmental and Educational Psychology, Faculty of Psychology, University of Vienna

Abstract

For human infants, the first years after birth are a period of intense exploration—getting to understand their own competencies in interaction with a complex physical and social environment. In contemporary neuroscience, the predictive-processing framework has been proposed as a general working principle of the human brain, the optimization of predictions about the consequences of one's own actions, and sensory inputs from the environment. However, the predictive-processing framework has rarely been applied to infancy research. We argue that a predictive-processing framework may provide a unifying perspective on several phenomena of infant development and learning that may seem unrelated at first sight. These phenomena include statistical learning principles, infants' motor and proprioceptive learning, and infants' basic understanding of their physical and social environment. We discuss how a predictive-processing perspective can advance the understanding of infants' early learning processes in theory, research, and application.

Keywords

cognition, infant development, neuroscience, perception, social cognition

The first years of life are a period of intense brain maturation and learning. It is a major challenge for young infants to acquire the capacities to navigate their complex environment. They have to make sense of the ends of their own actions through experiencing the proprioceptive effects of their body movement. At the same time, they have to acquire a basic understanding about the behavior of the physical and social entities in their environment. On the physiological level, this learning process relies on the continuous formation and pruning of connections in neuronal networks, which allows infants to interpret sensory information and translate their experience into appropriate behavioral responses in increasingly sophisticated ways. However, on the functional level, it remains an open theoretical and empirical question which basic principles underlie infants' early brain development and learning.

An influential theoretical account from cognitive neuroscience posits that the formation and refinement of predictive models is a major working principle of the human brain and that the main purpose of learning processes is to minimize prediction errors. The *predictive-processing* (PP) perspective originated from the basic computational problem that successful navigation in the environment relies on the organism's ability to optimize predictions about how one's own behavior will affect proprioceptive experiences (Helmholtz, 1867) and how social and physical entities in the outer world behave (e.g., Clark, 2013; Schubotz, 2015). Although the basic idea of the PP framework very closely resembles the challenges described for learning processes in young infants, researchers have only recently begun to investigate PP mechanisms in infancy (Emberson, Richards, & Aslin, 2015; Kayhan, Heil, et al., 2019; Kayhan, Hunnius, O'Reilly, & Bekkering, 2019; Kayhan, Meyer, O'Reily, Hunnius, & Bekkering, 2019; Kouider et al., 2015). The PP account has been related to several phenomena

Corresponding Author:
Moritz Köster, Freie Universität Berlin, Faculty of Education and Psychology, Habelschwerdter Allee 45, 14195 Berlin, Germany
E-mail: moritz.koester@fu-berlin.de

in developmental psychology, including mentalizing about own and others' internal bodily and mental states (Fotopoulou & Tsakiris, 2017; Palmer, Seth, & Hohwy, 2015), early language acquisition (Trainor, 2012), and autism spectrum disorder (e.g., Bolis & Schilbach, 2018; Lawson, Rees, & Friston, 2014; Pellicano & Burr, 2012; Sinha et al., 2014; Van de Cruys et al., 2014). However, it is still unclear how PP can inform our perspective on early learning processes more generally and whether it can be considered a unified framework to guide our understanding and research on cognitive development in the early infant years.

In this perspective article, we briefly summarize the core idea of the PP framework before highlighting how a PP perspective may provide a unifying account for several phenomena of infants' early learning. Phenomena that fit neatly into the PP explanatory framework include infants' statistical learning, motor learning, and proprioception as well as their emerging representations and expectations about the physical world. Finally, we discuss directions for future theoretical and empirical work.

Predictive Processing

The roots of the PP account go back to basic motor learning principles in the sense that the brain has to generate predictions of the organism's own motor outcomes (Helmholtz, 1867; for reviews on the theoretical origins, see Clark, 2013; Schubotz, 2015). This basic principle has recently been generalized and discussed as a basic working principle of the human brain (e.g., Friston, 2005, 2010). The principal problem the brain has to solve is the successful behavioral navigation in a dynamic physical and social environment. Every novel situation comes with uncertainties, which result from incomplete, noisy, and sparse sensory information. When perceiving and interacting with the environment, the brain has to deal with these uncertainties by making inferences and generating appropriate behavioral responses (Friston, 2005, 2010). That is, sensory inputs provide only highly incomplete information about a complex environment, and the sensory information available to the organism is highly variable and changes with its behavioral navigation. It is therefore essential for the brain to improve inferences on the basis of sensory inputs and minimizing prediction errors (Clark, 2013; Friston, 2005, 2010; Hohwy, 2007; Kwisthout, Bekkering, & van Rooij, 2017). As prediction errors are reduced, the accuracy of internal predictive models is increased (Friston, 2005, 2010). However, predictive models should not be misunderstood as rigid solutions to reoccurring situations but as a likelihood maximization of generative models based on former experiences. The adaptability of predictive models allows for adjustments to changing environments, including new social and physical contexts. Thus, the refinement of predictive models (i.e., the reduction of prediction errors) through learning allows the organism to interact with the environment in more and more competent ways.

A central idea of the PP framework is that the brain consists of lower and higher level areas, organized in a hierarchical system, and that the different levels continuously communicate with one another (Friston, 2005, 2010). Predictions are formed at every level of the hierarchy, from basic motor responses to higher reasoning, that is, from predictions about the outcomes of the organism's own actions (i.e., active inferences) to predictions about the physical world and actions and intentions of other agents (FeldmanHall & Shenhav, 2019; Kilner, Friston, & Frith, 2007; Koster-Hale & Saxe, 2013). Mismatches between what is predicted and what is perceived are critical incidences for learning. Prediction errors are sent back to higher levels in the hierarchy, where the prediction was made, to update existing predictions and thereby improve predictive models.

For instance, an infant may try to grab a ball at a particular location according to their internal predictive model of where the ball should be. If the ball has rolled away in the meantime, the prediction error (grabbing air instead of ball) is fed back to higher order areas involved in the prediction. The infant can now update their predictions on how the ball behaves (*perceptual inference*) or change the sensory input by moving their arm to where the ball is located (*active inference*) to minimize the prediction error. Yet, on a higher hierarchical level, the infant may understand that this ball was too fast for them to reach or even too far away, such that the infant would have to relocate their body to be able to reach the ball.

Predictive Processing Perspective on Infant Learning

In the following section, we highlight how the PP framework may offer a general account for several aspects of infant learning. We start with very basic learning processes, focusing on infants' statistical learning, and continue with infants' motor and proprioceptive learning as well as their early physical and social understanding. These phenomena fit neatly within the basic principles of the PP framework.

Statistical learning in infants

Bayesian accounts of brain functioning posit that the brain constantly computes the probability of events in the environment on the basis of incoming sensory information (Friston, 2005; Knill & Pouget, 2004). By acquiring statistical regularities of the environment, the organism forms probabilistic models of its environment. This makes statistical learning an essential mechanism in forming generative models of the environment on the basis of which predictions are made.

There is ample evidence in the literature showing that human infants generate probabilistic models to represent statistical regularities in their environment (Gopnik & Bonawitz, 2015; Ruffman, Taumoepeau, & Perkins, 2012). In their seminal work, Saffran, Aslin, and Newport (1996) showed that infants learn to segment artificial language on the basis of statistical information such as transitional probabilities between elements. In the habituation phase of their experiment, they presented 8-month-old infants pseudowords consisting of syllables that always followed each other (i.e., transitional probability of 1.0). In the test phase, they showed infants previously presented pseudowords as well as nonwords and used a novel syllable order. Infants looked longer at nonwords, which indicated that they learned the predictive structure of the pseudowords (for a recent review, see Saffran & Kirkham, 2018). Today, infants' statistical learning has been investigated extensively within different domains, including visual (Kirkham, Slemmer, & Johnson, 2002) and action domains (Monroy, Gerson, & Hunnius, 2017).

It seems reasonable to assume that the infant brain employs statistical learning principles to form predictions about basic contingencies in the environment, which is essential for the basic assumptions of the PP framework on the formation and optimization of predictive models. Below, we argue that the explanatory power of the PP framework is not restricted only to statistical learning. The PP framework provides a much broader theoretical framework on brain function and organization, including several aspects of behavioral navigation and concepts of the physical and social environment.

Infants' motor learning and proprioception

Early evidence for infants' behavioral adaptation to contingencies in their environment are their behavioral responses in associative-learning paradigms. For example, Ivkovich, Collins, Eckerman, Krasnegor, and Stanton (1999) first reported a conditioned eye-blink response when a tone was presented shortly before an air puff in 4- and 5-month-olds. Later, researchers showed that even 1-month-old infants can acquire a conditioned eye-blink response to social and nonsocial voice stimuli (Reeb-Sutherland et al., 2011). Thus, already young infants learn contingencies in their environment, in the basic sense of Hebbian learning (Hebb, 1949).

From early on, infants also learn about the consequences of their own actions in the environment. Rovee-Collier (1999) demonstrated that from 2 months onward, infants learned that a mobile connected to their feet moved in response to their own movement, as indicated by an increased kicking rate, in contrast to a condition in which the mobile was not connected to their feet. Infants even further increased their kicking rate when the mobile was disconnected, which indicates that they were trying to reproduce the effect (for a review, see Gerhardstein, Dickerson, Miller, & Hipp, 2012). Critically, computer simulations suggest that infants' increased kicking in the absence of an effect cannot be explained by simple action-effect binding, but it rather indicates that infants formed a causal link between the effect and their actions (Zaadnoordijk, Otworowska, Kwisthout, & Hunnius, 2018). The authors showed that a simulated infant robot, functioning on operant-conditioning principles without representing cause–effect relations, also increased its kicking rate when it was connected to the mobile but did not further increase its kicking after the mobile was disconnected. This finding suggests that infants represent their own actions (i.e., kicking) as the cause of effects in the world (i.e., moving mobile). This idea is in line with the idea that infants build up predictive models in interaction with their environment. When the effect ceases to appear (i.e., disconnected mobile), a prediction error occurs, and infants increase their effort (active inference) to adjust the world to their predictive models.

By the same token, infants' own interactions with the environment inform their increasingly sophisticated understanding of others' goal-directed actions (Stapel, Hunnius, Meyer, & Bekkering, 2016). When 3-month-old infants are provided with artificial grasping experiences through sticky mittens (gloves that facilitate infants' grasping experience), their sensitivity to an actor's goal increases (Sommerville, Woodward, & Needham, 2005). Furthermore, 6- to 10-month-old infants with high competencies to perform goal-directed grasping actions themselves are also better at predicting goal-directed grasping actions performed by others (Kanakogi & Itakura, 2011). Thus, infants seem to be able to use their own action-effect representations to make predictions about other agents' action goals, which suggests a close link between action generation and prediction (see also Keysers & Perrett, 2004).

This idea was formalized in a Bayesian model of infant imitation that suggested that infants' early motor behavior ("motor babbling") serves the generation of internal models and maps movements to consequences (Rao, Shon, & Meltzoff, 2007). In the form of forward models, these internal models can then be used to predict the consequences of one's own actions as well as the goals of another person's actions, which allows for (intention-based) imitation learning. Although the earliest forward models in motor development that likely begin to be formed in utero may not necessarily depend on prediction errors, later imitation learning seems to be, at least partly, fueled by infants' drive to minimize prediction error. For instance, infants are more likely to imitate a novel action when it is unexpected (e.g., because it is inefficient given the situational constraints of the model; Gergely, Bekkering, & Király, 2002). Observing such unexpected actions induces increased motor activation in the infant brain, which potentially reflects the updating of prior action predictions (Langeloh et al., 2018).

Evidence for changes in infants' active inferences about their own actions have been demonstrated explicitly with regard to their walking abilities, which emerge around the first birthday (Adolph & Tamis-LeMonda, 2014). Walking upright provides infants with a novel view on their physical world (Campos et al., 2000; Franchak, Kretch, Soska, & Adolph, 2011) that demands that they form novel predictive (proprioceptive and visual) models of the outcomes of their newly acquired actions. This has been shown for infants' wariness of heights. Just when infants begin to walk upright, they tend to readily walk and fall down a steep cliff of 90 cm (Adolph, 2000), whereas they would not, at the same age, go down this cliff in a crawling posture. However, when infants had experiences with upright walking in a baby walker, before they had begun to walk autonomously (i.e., making similar proprioceptive and visual experiences as in a walking posture), they could estimate the height of the cliff and avoid it (Dahl et al., 2013). Thus, infants adjust their predictive models about their novel abilities to interact with the environment on the basis of their prior sensory experiences (i.e., making sensory and active inferences; see also Anderson et al., 2013).

Given the uncertainties that come with infants' emerging capacities for walking, why do infants begin to walk in the first place? In developmental robotics, the imperative to learn about one's own motor capacities is often cast in terms of an intrinsic motivation (see also the concept of motor babbling; Barto, Mirolli, & Baldassarre, 2013; Friston et al., 2015, Kaplan & Oudeyer, 2007; Schmidhuber, 2010). We argue that intrinsic motivation is an integral aspect of PP that forms an important component of active inference. In active inference, agents form not only inferences about the causal structure that underlies their sensations but also inferences about specific actions that will minimize prediction error in the long run (Attias, 2003; Botvinick & Toussaint, 2012; Friston et al., 2015). That is, agents select those actions that resolve the greatest amount of uncertainty or, in short, that resolve uncertainty about "what would happen if I did that" (Schmidhuber, 2006). Resolving this uncertainty is exactly the intrinsic motivation or epistemic value associated with exploration of one's own motor capacities (Saegusa, Metta, Sandini, & Sakka, 2009). As we discuss below, these considerations link PP to infants' curiosity and exploration (i.e., choosing actions that afford the "opportunity for prediction error"), which is conceptually related to the notion of proximal development (Vygotsky, 1978). At this stage, it is sufficient to note that PP under active inference renders novelty seeking and responding to epistemic affordance a natural part of the way we forage for information in the service of self-modeling (and self-evidencing). These considerations closely correspond to the notion that young infants not only learn to move but also move to learn (e.g., Adolph, 2008).

These developmental phenomena in the motor domain neatly fit within the PP explanatory framework. Specifically, young infants already show appropriate motor responses to anticipated events in their environment and learn the consequences of their own actions on the basis of proprioceptive experiences in the interaction with their environment.

Infants' basic understanding about their environment

Infants' early learning about their social and physical environment is often discussed in terms of the formation of basic representations (e.g., Reid et al., 2009; Spelke, 1990; Wynn, 1992). Infants focus their attention selectively on novel events and objects with which they are not familiar (novelty preference; Fantz, 1965). That is, infants lose interest in perceptual stimuli that they have repeatedly encountered (*habituation*), and their attention revives for novel stimuli (*dishabituation*). Traditionally, infants' interest in novel events and objects has been described in terms of a *comparator model* (Sokolov, 1963, 1990): When orienting toward a stimulus, the infant compares the sensory information with an existing neuronal representation. If the current stimulus deviates from the existing representation, an orienting response leads to increased attention and the formation or update of a neuronal representation of the respective stimulus.

Critically, it has recently been acknowledged in cognitive science that there is no one-on-one correspondence between sensory inputs and mental representations. For instance, the sensory information of a stimulus changes dramatically after movements in relation to the specific stimulus, which requires the organism to understand the contingencies between behavioral responses and changes in sensory input (O'Regan & Noë, 2001). Thus, the main purpose of processing sensory information is not the mere representation of the external world but the generation of appropriate behavioral responses by making inferences on the consequences of behavioral responses for sensory inputs (sensorimotor contingency perspective; Engel, Maye, Kurthen, & König, 2013; O'Regan & Noë, 2001). In this sense, in contrast to the acquisition of representations, from a PP perspective, infants' learning may be conceptualized as the formation and refinement of predictive models about animate and physical entities in relation to the infant's own body movements and actions.

Infants' emerging understanding about the physical and social world is commonly investigated by using violation-of-expectation (VOE) paradigms. In VOE paradigms, infants' orienting response (Sokolov, 1963, 1990) is taken as an indicator of infants' basic concepts about the environment. Unexpected events that violate physical or social rules lead to differential responses in infants' gaze behavior (e.g., Wynn, 1992), pupil dilation (Gredebäck & Melinder, 2010), and event-related brain responses (Berger, Tzur, & Posner, 2006; Köster, Langeloh, & Hoehl, 2019; Langeloh et al., 2018; Reid et al., 2009). For example, infants detect impossible physical events (e.g., a ball rolling through a wall; Spelke, Breinlinger, Macomber, & Jacobson, 1992), changes in numbers (e.g., changes in the number of toys behind an occluder; Simon, Hespos, & Rochat, 1995; Wynn, 1992), or irrational human actions (e.g., a pretzel that is put toward the ear instead of the mouth; Reid et al., 2009). The PP perspective offers a plausible explanation for infants' VOE responses across social and physical knowledge domains. VOE responses indicate infants' processing of prediction errors, which require them to refine prior predictions (i.e., update their predictive models) or, put more simply, learning. It was formerly emphasized that unexpected events provide infants with novel opportunities to learn (Baillargeon, 2004; Leslie, 2004; Stahl & Feigenson, 2019). In support of this view, a recent study of 11-month-old infants demonstrated an increase in subsequent exploration and hypothesis-testing behavior by the infants for objects that behaved in an unexpected way (Stahl & Feigenson, 2015). Thus, infants actively seek to reduce their uncertainties, which may be particularly important for objects that do not comply with their existing predictive models.

In sum, the PP account provides a framework that is compatible with several well-known phenomena regarding infants' processing of novel and unexpected information in their environment. That is, beyond the comparator model, it also holds as an explanatory framework for the processing and imitation of unexpected actions and the exploration behavior toward objects that behaved in an unexpected way. Furthermore, the PP perspective on infants' early understanding of their environment highlights that these phenomena are largely compatible with the paradigm shift in the cognitive sciences, away from a representation-based account toward an enacted account of human cognition (Engel et al., 2013; O'Regan & Noë, 2001).

Future Perspectives

By sorting different developmental phenomena roughly along their level of complexity, we have pointed out how infants learn, from very early in life, to predict the effects of their own actions and basic regularities in their environment and how they use novel experiences to adjust their predictive models. We will now discuss how a PP perspective can advance our understanding of infants' early social learning and its implications for future research.

Mastering uncertainties: A motivation to learn and a need for structure

A critical aspect of the PP perspective is that the organism seeks to form predictive models to handle uncertainties. Friston (2010) based this consideration on the free energy principle and stated that organisms seek to reduce entropy. This idea has intriguing implications for infants' motivation to explore and learn from their environment and likewise for their preference for behavioral regularities. From a PP perspective, the motivation to learn, to reduce prediction errors, may be a self-sufficient process that explains infants' curiosity (i.e., their intrinsic motivation to explore; Twomey & Westermann, 2018) and interest in novel objects, events, and activities (Stahl & Feigenson, 2019; see also the section on infants' motor learning and proprioception). The emphasis is on building up and optimizing internal models (i.e., the importance of novelty seeking and the epistemic value in early exploratory behavior). In this instance, exactly the same principles that underlie motor exploration may also drive infants' general exploratory tendencies—resolving uncertainty about the way that the world works (see also Kayhan & Kwisthout, 2017, p. 10).

For young infants, interacting with social agents in particular comes with many uncertainties given that the thoughts and intentions of other individuals are largely hidden and people often act differently in similar situations (FeldmanHall & Shenhav, 2019). FeldmanHall and Shenhav (2019) argued that specific strategies have evolved to improve predictions in a social context, including impression formation and perspective taking. Furthermore, Bayesian models of social learning, including epistemic trust and imitation, take advantage of the fact that probabilistic models enable the integration of multiple kinds of data (i.e., statistical and social information; e.g., Gopnik & Bonawitz, 2015; Rao et al., 2007; Shafto & Goodman, 2008). We would like to add to these considerations that human social norms, behavioral codes within social groups that regulate human social interactions, may have evolved, to some extent, to reduce uncertainties in the social world (for a similar argument, see Veissière, Constant, Ramstead, Friston, & Kirmayer, 2019). From early on, infants seek to understand (Köster & Hepach, 2019), comply with (Haun, Rekers, & Tomasello, 2014), and enforce social norms (Rakoczy & Schmidt, 2013). In a similar vein, it has been argued that social learning, resulting in decreased prediction errors in complex social exchanges, may thus be inherently rewarding and, at the same time, help others to reduce their prediction errors (de Bruin & Michael, 2018). Thus, a PP perspective has theoretical implications for infants' motivation to learn about their complex social environments and their preference for structured environments.

Limitations of the predictive processing account

Despite its success, the PP framework has gotten its share of criticism. One of the main criticisms of the framework regards its claim to be a unifying theory of the brain (Colombo & Wright, 2017; Klein, 2018). Colombo and Wright (2017) argued that the brain is a complex interplay between multiple systems, which renders any grand unifying hypothesis unjustified. Instead, neuroscientific progress is suggested to rely on experimental and theoretical work trying to answer different smaller questions simultaneously. Another major criticism of the PP framework regards its testability. For example, Kogo and Trengove (2015) have argued that the PP framework does not specify how the more fine-grained neuronal mechanisms underlying the basic computational principles, such as error computation and minimization, are implemented in the neurophysiology of the brain.

Applying these critiques to infancy research, central questions are (a) how a unifying perspective on several learning phenomena may advance our overall understanding of infant brain development and learning and (b) which specific testable hypotheses we can derive from a PP perspective on infant learning.

Understanding predictive processes in the infant brain

It is, in our view, remarkable how a PP framework is compatible with and unifies central phenomena in infants' early learning at very different levels, from statistical learning to motor development and social learning. However, research on PP in the infant brain is in its infancy. Although there is accumulating evidence that the adult brain might be working on the basis of the principles of the PP framework (e.g., Egner, Monti, & Summerfield, 2010; Wacongne et al., 2011), our knowledge on the predictive nature of the infant brain is sparse (but for recent research on PP in infants, see Emberson et al., 2015; Kayhan, Heil, et al., 2019; Kayhan, Hunnius, et al., 2019; Kayhan, Meyer, et al., 2019; Kouider et al., 2015). These studies provided initial evidence that the infant brain is already capable of forming predictions on the basis of prior knowledge and physiologically responds to violations of these predictions.

Oakes and Rakison (2019) have pointed out that developmental attainments built on each other and that early developmental changes in a specific domain lay the ground for future developments. This idea corresponds closely to a cornerstone of Bayesian learning: Each time new evidence is received, the prior (predictive model) is updated to a new posterior (adjusted model), which becomes a new prior for making inferences when next evidence is observed. To give an example: Priors of the visual environment in relation to the body are updated when the infant learns to stand up or is put in a walker. The adjusted models then help the infant navigate the environment when moving on two feet (Dahl et al., 2013). Furthermore, according to Oakes and Rakison (2019), developmental cascades originate in the structure and development of the nervous system. This notion is highly interesting regarding the hierarchical structure proposed by the PP framework. Predictive models are formed at each level of the hierarchy and in different domains, and the interplay between different levels and domains is not well understood. For example, many physical and social concepts emerge before infants can put them into words (e,g, Bergelson & Swingley, 2012) or action (e.g., Köster, Itakura, Omori, & Kärtner, 2019; see also Köster & Kärtner, 2019), which suggests that these cognitive capacities develop independently from specific motor abilities. For future research, it will be intriguing to understand the development at different levels of the hierarchy and interdependency between different hierarchical levels.

From a developmental cognitive neuroscience view, crucial questions are at which ages specific predictive processes develop and how these relate to structural and neuro-computational changes in the maturing infant brain. Is the infant brain indeed a "prediction machine" (cf. Clark, 2013), constantly generating predictions about own action outcomes and the social and physical world? How do infants use prediction errors to update predictive models? Which are the neural markers of the formation and refinement of predictive models in the infant brain? Which are the optimal learning conditions, and how big should prediction errors be to facilitate the updating of internal models? How do these computational foundations interact with basic human needs and motives? Answering these questions will be essential to better understand how the PP framework can inform our conception of infant learning and early cognitive development.

From an applied perspective, understanding PP in the context of infant brain development and learning may guide the design of child care and learning environments. For example, these environments could be designed to stimulate infant learning by offering an optimal balance between predictability (so as not to overwhelm infants) and opportunity for prediction error such that infants can build up and fine-tune more and more complex and precise internal models of the world (closely resembling the Vygotskian idea of proximal development; Vygotsky, 1978). Moreover, it has recently been argued that a predictive processing account as a general framework for early human development may have important implications for developmental robotics (Nagai, 2019).

Overall, although the PP framework is a prevailing account of human brain functioning and organization, it has, to date, been applied only to specific domains of early human development and learning. We have highlighted how the PP account may provide a useful and unified theoretical framework to understand a range of developmental phenomena in early infancy. For a better understanding of PP in the infant brain, it will be crucial to much better understand basic computational principles of PP at early developmental stages and show their implications for developmental trajectories across several domains of infant learning. This work would make a strong case for the PP as a general framework for research and theory on infant development and learning.

Transparency

Action Editor: Jennifer Wiley
Editor: Laura A. King
Declaration of Conflicting Interests
The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

ORCID iDs

Moritz Köster https://orcid.org/0000-0002-9951-7741
Ezgi Kayhan https://orcid.org/0000-0002-1725-220X
Miriam Langeloh https://orcid.org/0000-0003-3250-8602

References

Adolph, K. E. (2000). Specificity of learning: Why infants fall over a veritable cliff. *Psychological Science*, *11*, 290–295. doi:10.1111/1467-9280.00258

Adolph, K. E. (2008). Learning to move. *Current Directions in Psychological Science*, *17*, 213–218. doi:10.1111/j.1467-8721.2008.00577.x

Adolph, K. E., & Tamis-LeMonda, C. S. (2014). The costs and benefits of development: The transition from crawling to walking. *Child Development Perspectives*, *8*, 187–192. doi:10.1111/cdep.12085

Anderson, D. I., Campos, J. J., Witherington, D. C., Dahl, A., Rivera, M., He, M., . . . Barbu-Roth, M. (2013). The role of locomotion in psychological development. *Frontiers in Psychology*, *4*, Article 440. doi:10.3389/fpsyg.2013.00440

Attias, H. (2003, January). *Planning by probabilistic inference*. In C. M. Bishop & B. J. Frey (Eds.), Paper presented at the Ninth International Workshop on Artificial Intelligence and Statistics (AISTATS 2003), Key West, FL.

Baillargeon, R. (2004). Infants' physical world. *Current Directions in Psychological Science*, *13*, 89–94. doi:10.1111/j.0963-7214.2004.00281.x

Barto, A., Mirolli, M., & Baldassarre, G. (2013). Novelty or surprise? *Frontiers in Psychology*, *4*, Article 907. doi:10.3389/fpsyg.2013.00907

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences, USA*, *109*, 3253–3258. doi:10.1073/pnas.1113380109

Berger, A., Tzur, G., & Posner, M. I. (2006). Infant brains detect arithmetic errors. *Proceedings of the National Academy of Sciences, USA*, *103*, 12649–12653. doi:10.1073/pnas.0605350103

Bolis, D., & Schilbach, L. (2018). Developmental cognitive neuroscience observing and participating in social interactions: Action perception and action control across the autistic spectrum. *Developmental Cognitive Neuroscience*, *29*, 168–175. doi:10.1016/j.dcn.2017.01.009

Botvinick, M., & Toussaint, M. (2012). Planning as inference. *Trends in Cognitive Sciences*, *16*(10), 485–488.

Campos, J. J., Anderson, D. I., Barbu-Roth, M. A., Hubbard, E. M., Hertenstein, M. J., & Witherington, D. (2000). Travel broadens the mind. *Infancy*, *1*, 149–219. doi:10.1207/S15327078IN0102_1

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204.

Colombo, M., & Wright, C. (2017). Explanatory pluralism: An unrewarding prediction error for free energy theorists. *Brain and Cognition*, *112*, 3–12.

Dahl, A., Campos, J. J., Anderson, D. I., Uchiyama, I., Witherington, D. C., Ueno, M., . . . Barbu-Roth, M. (2013). The epigenesis of wariness of heights. *Psychological Science*, *24*, 1361–1367. doi:10.1177/0956797613476047

de Bruin, L., & Michael, J. (2018). Prediction error minimization as a framework for social cognition research. *Erkenntnis*. doi:10.1007/s10670-018-0090-9

Egner, T., Monti, J. M., & Summerfield, C. (2010). Expectation and surprise determine neural population responses in the ventral visual stream. *Journal of Neuroscience*, *30*, 16601–16608. doi:10.1523/jneurosci.2770-10.2010

Emberson, L. L., Richards, J. E., & Aslin, R. N. (2015). Top-down modulation in the infant brain: Learning-induced expectations rapidly affect the sensory cortex at 6 months. *Proceedings of the National Academy of Sciences, USA*, *112*, 9585–9590. doi:10.1073/pnas.1510343112

Engel, A. K., Maye, A., Kurthen, M., & König, P. (2013). Where's the action? The pragmatic turn in cognitive science. *Trends in Cognitive Sciences*, *17*, 202–209. doi:10.1016/J.TICS.2013.03.006

Fantz, R. L. (1965). Visual perception from birth as shown by pattern selectivity. *Annals of the New York Academy of Sciences*, *118*, 793–814. doi:10.1111/j.1749-6632.1965.tb40152.x

FeldmanHall, O., & Shenhav, A. (2019). Resolving uncertainty in a social world. *Nature Human Behaviour*, *3*, 426–435. doi:10.1038/s41562-019-0590-x

Fotopoulou, A., & Tsakiris, M. (2017). Mentalizing homeostasis: The social origins of interoceptive inference. *Neuropsychoanalysis*, *19*(1), 3–28.

Franchak, J. M., Kretch, K. S., Soska, K. C., & Adolph, K. E. (2011). Head-mounted eye tracking: A new method to describe infant looking. *Child Development*, *82*, 1738–1750. doi:10.1111/j.1467-8624.2011.01670.x

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*, 815–836. doi:10.1098/rstb.2005.1622

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

Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., & Pezzulo, G. (2015). Active inference and epistemic value. *Cognitive Neuroscience*, *6*(4), 187–214.

Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature*, *415*(6873), 755–755.

Gerhardstein, P., Dickerson, K., Miller, S., & Hipp, D. (2012). Early operant learning is unaffected by socio-economic status and other demographic factors: A meta-analysis. *Infant Behavior and Development*, *35*, 472–478. doi:10.1016/j.infbeh.2012.02.005

Gopnik, A., & Bonawitz, E. (2015). Bayesian models of child development. *WIREs Cognitive Science*, *6*(2), 75–86. doi:10.1002/wcs.1330

Gredebäck, G., & Melinder, A. (2010). Infants' understanding of everyday social interactions: A dual process account. *Cognition*, *114*, 197–206. doi:10.1016/j.cognition.2009.09.004

Haun, D. B. M., Rekers, Y., & Tomasello, M. (2014). Children conform to the behavior of peers; other great apes stick with what they know. *Psychological Science*, *25*, 2160–2167. doi:10.1177/0956797614553235

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. New York, NY: Wiley.

Helmholtz, H. V. (1867). *Handbuch der physiologischen Optik* [Manual of physiological optics]. Leipzig, Germany: Leopold Voss.

Hohwy, J. (2007). The sense of self in the phenomenology of agency and perception. *Psyche*, *13*, 1–20. Retrieved from http://journalpsyche.org/files/0xab11.pdf

Ivkovich, D., Collins, K. L., Eckerman, C. O., Krasnegor, N. A., & Stanton, M. E. (1999). Classical delay eyeblink conditioning in 4- and 5-month-old human infants. *Psychological Science*, *10*, 4–8. doi:10.1111/1467-9280.00097

Kanakogi, Y., & Itakura, S. (2011). Developmental correspondence between action prediction and motor ability in early infancy. *Nature Communications*, *2*(1), Article 341. doi:10.1038/ncomms1342

Kaplan, F., & Oudeyer, P. Y. (2007). In search of the neural circuits of intrinsic motivation. *Frontiers in Neuroscience*, *1*(1), 225–236. doi:10.3389/neuro.01.1.1.017.2007

Kayhan, E., Heil, L., Kwisthout, J., van Rooij, I., Hunnius, S., & Bekkering, H. (2019). Young children integrate current observations, priors and agent information to predict others' actions. *PLOS ONE*, *14*(5), Article e0200976. doi:10.1371/journal.pone.0200976

Kayhan, E., Hunnius, S., O'Reilly, J. X., & Bekkering, H. (2019). Infants differentially update their internal models of a dynamic environment. *Cognition*, *186*, 139–146. doi:10.1016/J.COGNITION.2019.02.004

Kayhan, E., & Kwisthout, J. (2017). Predictive processing in development. *CDS Newsletter: The Newsletter of the Technical Committee on Cognitive and Developmental Systems*, *141*(1), 10.

Kayhan, E., Meyer, M., O'Reilly, J. X., Hunnius, S., & Bekkering, H. (2019). Nine-month-old infants update their predictive models of a changing environment. *Developmental Cognitive Neuroscience*, *38*, Article 100680. doi:10.1016/j.dcn.2019.100680

Keysers, C., & Perrett, D. I. (2004). Demystifying social cognition: A Hebbian perspective. *Trends in Cognitive Sciences*, *8*, 501–507. doi:10.1016/j.tics.2004.09.005

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: An account of the mirror neuron system. *Cognitive Processing*, *8*, 159–166. doi:10.1007/s10339-007-0170-2

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*(2), B35–B42. doi:10.1016/S0010-0277(02)00004-5

Klein, S. B. (2018). *Learning: principles and applications* (8th ed.). Thousand Oaks, CA: SAGE.

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

Kogo, N., & Trengove, C. (2015). Is predictive coding theory articulated enough to be testable? *Frontiers in Computational Neuroscience*, *9*, Article 111. doi:10.3389/fncom.2015.00111

Köster, M., & Hepach, R. (2019). *Normative expectations in preverbal infants*. Manuscript submitted for publication.

Köster, M., Itakura, S., Omori, M., & Kärtner, J. (2019). From understanding others' needs to prosocial action: Motor and social abilities promote infants' helping. *Developmental Science*, *22*(6), Article e12804. doi:10.1111/desc.12804

Köster, M., & Kärtner, J. (2019). Why do infants help? A simple action reveals a complex phenomenon. *Developmental Review*, *51*, 175–187.

Köster, M., Langeloh, M., & Hoehl, S. (2019). Visually entrained theta oscillations increase for unexpected events in the infant brain. *Psychological Science*, *30*, 1656–1663.

Koster-Hale, J., & Saxe, R. (2013). Theory of mind: A neural prediction problem. *Neuron*, *79*(5), 836–848.

Kouider, S., Long, B., Le Stanc, L., Charron, S., Fievet, A. C., Barbosa, L. S., & Gelskov, S. V. (2015). Neural dynamics of prediction and surprise in infants. *Nature Communications*, *6*, Article 8537. doi:10.1038/ncomms9537

Kwisthout, J., Bekkering, H., & van Rooij, I. (2017). To be precise, the details don't matter: On predictive processing, precision, and level of detail of predictions. *Brain and Cognition*, *112*, 84–91.

Langeloh, M., Buttelmann, D., Matthes, D., Grassmann, S., Pauen, S., & Hoehl, S. (2018). Reduced Mu power in response to unusual actions is context-dependent in 1-year-olds. *Frontiers in Psychology*, *9*, Article 36. doi:10.3389/fpsyg.2018.00036

Lawson, R. P., Rees, G., & Friston, K. J. (2014). An aberrant precision account of autism. *Frontiers in Human Neuroscience*, *8*, Article 302. doi:10.3389/fnhum.2014.00302

Leslie, A. (2004). *Who's for learning?* Retrieved from https://psycnet.apa.org/record/2004-17725-004

Monroy, C., Gerson, S., & Hunnius, S. (2017). Infants' motor proficiency and statistical learning for actions. *Frontiers in Psychology*, *8*, Article 2174. doi:10.3389/fpsyg.2017.02174.

Nagai, Y. (2019). Predictive learning: Its key role in early cognitive development. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*. doi:10.1098/rstb.2018.0030

Oakes, L. M., & Rakison, D. H. (2019). *Developmental cascades: Building the infant mind*. New York, NY: Oxford University Press.

O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral & Brain Sciences*, *24*, 939–973. doi:10.1017/S0140525X01000115

Palmer, C. J., Seth, A. K., & Hohwy, J. (2015). The felt presence of other minds: Predictive processing, counterfactual predictions, and mentalising in autism. *Consciousness and Cognition*, *36*, 376–389.

Pellicano, E., & Burr, D. (2012). When the world becomes 'too real': A Bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, *16*(10), 504–510.

Rakoczy, H., & Schmidt, M. F. H. (2013). The early ontogeny of social norms. *Child Development Perspectives*, *7*, 17–21. doi:10.1111/cdep.12010

Rao, R. P., Shon, A. P., & Meltzoff, A. N. (2007). A Bayesian model of imitation in infants and robots. In C. Nehaniv & K. Dautenhahn (Eds.), *Imitation and social learning in robots, humans, and animals* (pp. 217–247). Cambridge, England: Cambridge University Press.

Reeb-Sutherland, B. C., Fifer, W. P., Byrd, D. L., Hammock, E. A. D., Levitt, P., & Fox, N. A. (2011). One-month-old human infants learn about the social world while they sleep. *Developmental Science*, *14*, 1134–1141. doi:10.1111/j.1467-7687.2011.01062.x

Reid, V. M., Hoehl, S., Grigutsch, M., Groendahl, A., Parise, E., & Striano, T. (2009). The neural correlates of infant and adult goal prediction: Evidence for semantic processing systems. *Developmental Psychology*, *45*, 620–629. doi:10.1037/a0015209

Rovee-Collier, C. (1999). The development of infant memory. *Current Directions in Psychological Science*, *8*, 80–85. doi:10.1111/1467-8721.00019

Ruffman, T., Taumoepeau, M., & Perkins, C. (2012). Statistical learning as a basis for social understanding in children. *British Journal of Developmental Psychology*, *30*, 87–104. doi:10.1111/j.2044-835X.2011.02045.x

Saegusa, R., Metta, G., Sandini, G., & Sakka, S. (2009). Active motor babbling for sensorimotor learning. In *2008 IEEE International Conference on Robotics and Biomimetics* (pp. 794-799). doi:10.1109/ROBIO.2009.4913101

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology*, *69*, 181–203. doi:10.1146/annurev-psych-122216-011805

Schmidhuber, J. (2006). Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts. *Connection Science*, *18*(2), 173–187.

Schmidhuber, J. (2010). Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Transactions on Autonomous Mental Development*, *2*(3), 230–247. doi:10.1109/TAMD.2010.2056368

Schubotz, R. I. (2015). Prediction and expectation. In A. W. Toga (Ed.), *Brain mapping: An encyclopedic reference* (Vol. 3, pp. 295-302). doi:10.1016/B978-0-12-397025-1.00247-5

Shafto, P., & Goodman, N. (2008). Teaching games: Statistical sampling assumptions for learning in pedagogical situations. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th annual conference of the Cognitive Science Society* (pp. 1632–1637). Retrieved from http://csjarchive.cogsci.rpi.edu/proceedings/2008/pdfs/p1632.pdf

Simon, T. J., Hespos, S. J., & Rochat, P. (1995). Do infants understand simple arithmetic? A replication of Wynn (1992). *Cognitive Development*, *10*, 253–269. doi:10.1016/0885-2014(95)90011-X

Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., & Held, R. M. (2014). Autism as a disorder of prediction. *Proceedings of the National Academy of Sciences, USA*, *111*, 15220–15225. doi:10.1073/pnas.1416797111

Sokolov, E. N. (1963). *Perception and the conditioned reflex* (S. W. Waydenfeld, Trans.). Oxford, England: Pergamon Press.

Sokolov, E. N. (1990). The orienting response, and future directions of its development. *The Pavlovian Journal of Biological Science*, *25*, 142–150. https://doi.org/10.1007/BF02974268

Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition*, *96*, 1–11. doi:10.1016/j.cognition.2004.07.004

Spelke, E. S. (1990). Principles of object perception. *Cognitive Science*, *14*, 29–56. doi:10.1016/0364-0213(90)90025-R

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge part 2. *Psychological Review*, *99*, 605–632. Retrieved from https://psycnet.apa.org/record/1993-05134-001

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science*, *348*(6230), 91–94. doi:10.1126/science.aaa3799

Stahl, A. E., & Feigenson, L. (2019). Violations of core knowledge shape early learning. *Topics in Cognitive Science*, *11*, 136–153. doi:10.1111/tops.12389

Stapel, J. C., Hunnius, S., Meyer, M., & Bekkering, H. (2016). Motor system contribution to action prediction: Temporal accuracy depends on motor experience. *Cognition*, *148*, 71–78. doi:10.1016/j.cognition.2015.12.007

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. doi:10.1016/J.IJPSYCHO.2011.12.008

Twomey, K. E., & Westermann, G. (2018). Curiosity-based learning in infants: A neurocomputational approach. *Developmental Science*, *21*(4), Article e12629. doi:10.1111/desc.12629

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: Predictive coding in autism. *Psychological Review*, *121*(4), 649–675.

Veissière, S. P. L., Constant, A., Ramstead, M. J. D., Friston, K. J., & Kirmayer, L. J. (2019). Thinking through other minds: A variational approach to cognition and culture. *Behavioral & Brain Sciences*, 1–97. doi:10.1017/S0140525X19001213

Vygotsky, L. (1978). Interaction between learning and development. *Readings on the Development of Children*, *23*(3), 34–41.

Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proceedings of the National Academy of Sciences, USA*, *108*, 20754–20759. doi:10.1073/pnas.1117807108

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, *358*, 749–750.

Zaadnoordijk, L., Otworowska, M., Kwisthout, J., & Hunnius, S. (2018). Can infants' sense of agency be found in their behavior? Insights from babybot simulations of the mobile-paradigm. *Cognition*, *181*, 58–64. doi:10.1016/j.cognition.2018.07.006