

**Danaja Rutar**

## Developing higher cognition through predictive processing

**Abstract:** Predictive processing is an influential theoretical framework in cognitive science that promises to provide a unifying account of all perception, action and cognition. Whilst it has already successfully explained many facets of perception and action, crucially, it is lacking a coherent theoretical account of cognitive development and higher cognition and, even more ambitiously, the *development of higher order thought*. In this paper, I investigate to what extent predictive processing in its current form is even suitable for explaining development and higher thought. I suggest that as things stand, predictive processing cannot sufficiently explain the development of higher order as the internal models it proposes lack two crucial ingredients that characterise higher thought: compositionality and generativity. I propose how these two core features can be addressed within predictive processing. Lastly, I elaborate on the reciprocal relationship between pedagogical practices and predictive processing: pedagogical practices are crucial for model building and predictive processing offers a theoretical framework for understanding psychological mechanisms involved in pedagogical settings.

**Keywords:** predictive processing, development of higher thought, compositionality, generativity, pedagogical implications

UDC: 159.95

Scientific article

## Introduction

Predictive processing (PP from now onwards) is a key theoretical framework in cognitive science that promises to explain the entirety of cognition, perception and action with one set of core mechanisms (Friston 2009, 2010; Clark 2013, 2015). PP is built upon Bayesian theories of the brain, and as such, it postulates that the brain is a »probabilistic machine« that is continuously predicting the world and updating its generative model based on the discrepancy between what it predicts and what it perceives (Clark 2013, 2015). PP has been successfully applied to many different aspects of cognitive functioning, such as perception (Kok et al. 2013; Muckli et al. 2015; Summerfield and de Lange 2014), action understanding (Urgen and Miller 2015), planning and navigation (Kaplan and Friston 2018), communication (Friston and Penny 2011), learning (Friston et al. 2017; Kwisthout et al. 2017; Rutar et al. 2023; Rutar et al. 2022; Smith et al. 2020) and even mentalising (Koster-Hale and Saxe 2013; Heil et al. 2019). Whilst PP has been applied to a great many domains, the majority of success has been achieved with regard to perceptual (as opposed to higher cognitive) processes in adults.

If PP claims to provide a unifying theory for all aspects of human functioning – perception, action and cognition – accounting for the topics that are poorly addressed in PP, such as development and higher cognition, should be the priority of PP research. Investigation of higher cognition might prove to be specifically thorny, however, as PP was initially designed as a data compression method for studying information processing on the retina (Srinivasan et al. 1982). PP has been primarily used to try to explain low-level perceptual processes, and it is not trivial how and if it can scale up to higher cognition. In particular, it is not clear whether the kinds of internal models that underlie perceptual processes can sufficiently account for higher cognitive processes (Clark 2013, p. 201): »What [...] the local approximations to Bayesian reasoning look like as I depart further and further from the safe shores of basic perception and motor control? What new forms of representation are then required, and how do they behave in the context of the hierarchical predictive coding regime?«

As for motivating the investigation of the developmental origins of PP, any

account of cognition must be developable; otherwise, claims about adult cognition have weak explanatory value. Namely, the plausibility of adult capacities is rather limited if no explanation is provided as to how such capacities came to be (Ward et al. under review). This holds truer still if a theory aims to provide a unifying account of human functioning, such as that of PP.

The PP community is slowly starting to recognise the importance of addressing development and higher cognition. For example, in recent years, evidence that predictive machinery is in place in young children has started to pile up (Kayhan et al. 2019; Kayhan et al. 2019; Nagai 2019; Zhang et al. 2019), claims have been made that PP could be compatible with existing developmental findings (Köster et al. 2020) and first attempts at delineating developmental PP have begun to emerge (Ward et al. under review). Similarly, research considering higher aspects of cognition have started to take off (Heil et al. 2019; Neacsu et al. 2022; Rutar et al. 2022; Rutar et al. 2023; Smith et al. 2020).

However, what we are still lacking is a mechanistic account of higher-order cognition *and* a theory of development or, even more ambitiously, a theory of the development of higher cognition. I will treat development and higher cognition (from the PP perspective) in synchrony. In what follows, I articulate the requirements that internal models and the learning mechanisms in PP will need to fulfil to allow for the development of higher thought. In fleshing out said requirements, I will consider two questions: What is the nature of internal models that support higher order thought? And what learning mechanisms underlie the development and refinement of these internal models? I start by presenting the core tenets of the PP theory and then address each of the questions in turn: in relation to the first question, I discuss how the models that support higher cognition are *compositional* (Fodor 1975; Fodor and Pylyshyn 1988) and in relation to the second question, I discuss *generativity* as a crucial aspect of internal models that allows for further development and learning (Fodor 1975; Fodor and Pylyshyn 1988). Each of these two characteristics, compositionality and generativity, are discussed from the point of view of PP as well. Lastly, I explore how pedagogical practices are detrimental for model building and how PP can offer a useful theoretical framework for understanding psychological processes involved in a pedagogical setting.

## What is PP?

### *Generative models compute predictions, hypotheses and prediction error*

According to PP, the brain functions as a hierarchically organised generative model that computes increasingly abstract hypotheses about the world. More specifically, at each level, the brain computes *hypotheses* about the hidden causes of the input from the layer below (Clark 2015; Friston 2010; Kwisthout et al. 2017; Friston et al. 2016). The generative model also continuously makes *predictions* about observable *outcomes* in terms of discrete events (Clark 2013; Friston et al. 2016; Hohwy 2013; Kwisthout et al. 2017). More informally, the model predicts

what it is going to experience next. For example, in the context of playing Don't get angry! game predictions will pertain to the outcome of a dice roll (numbers from one to six), and when talking with a friend, predictions will pertain to the friend's emotions (happy, sad, angry). Across the entirety of the generative model, predictions from the layers above and the sensory input from the layers below are compared – the difference between the two is computed and is called *prediction error* (Clark 2013, 2015; Friston and Kiebel 2009; Hohwy 2013). According to PP, the ultimate goal of any living system is to *minimise* prediction error in the long run (Friston 2009, 2010). Lower layers of the model are involved in processing concrete perceptual details, whereas the higher layers code for more abstract aspects of events that can be more temporally and spatially remote (Clark 2013; Hohwy 2013). For example, predictions at lower layers will contain information about patterns of light on the retina, and complementary predictions at higher layers will contain information about which object might have caused those patterns of light on the retina.

### *Minimising prediction error*

There are two ways in which a cognitive system can minimise prediction error: through perceptual inference and active inference (Kwisthout et al. 2017). *Perceptual inference* or *Bayesian model updating* pertains to changing the current probability distribution over the hypotheses at different layers of the hierarchy until there is no prediction error left. Bayesian model updating is an iterative process that happens every time the brain receives new sensory evidence (Bayes 1763/1958; Oaksford and Chater 2009; O'Reilly et al. 2013). With every new piece of evidence, the hypothesis changes slightly, which is reflected in the new estimate of the mean and variance of the hypothesis. This process leads to more accurate predictions in the future (Clark 2013; Friston 2010; Kwisthout et al. 2017). For example, suppose that you meet a new friend and after a period of meeting up with them, you start noticing a pattern: your friend is always approximately 10 minutes late for your appointment. What happens at the level of the model is that the probability distribution over the hypothesis that predicts your friend coming at the agreed-upon time is slowly changing. Over time, the hypothesis that your friend will be on time becomes less and less likely; thus, when your friend is yet again 10 minutes late, this occurrence generates less prediction error. Alternatively, prediction error can be minimised through *active inference*, which means sampling new sensory evidence to bring it more in line with predictions, which also leads to more accurate predictions in the future (Friston 2010; Kwisthout et al. 2017; Friston et al. 2016). For example, a dancer makes a jump and accidentally lands on a co-dancer. The failed jump generates a proprioceptive prediction error between where she expected to land and where she ended up landing. In light of this prediction error, she changes her position to where she intended to land, bringing her position closer to the predicted one.

## PP account of higher-order cognition

Having presented the core features of PP, I now move to the core part of the paper. In the following, I will investigate to what extent PP is as it stands able to explain the development of higher cognition. In addressing this question, I will investigate two sub-questions: What is the nature of internal models that support higher-order thought? And what learning mechanisms allow for the development and refinement of these internal models? Discussing models and learning requirements in relation to these questions will provide the basis for further evaluation of the extent to which PP satisfies these requirements.

*What is the nature of internal models that support higher-order thought?*

Higher thought is compositional

A core feature of the human (both adult and children) mind is that it is capable of thinking about anything, imagining things that it has never seen, reimagining events that have happened long ago, solving novel problems, generating new beliefs, coming up with new concepts and even new languages, to mention just a few examples of this remarkable ability. One common denominator of these higher cognitive processes is that they are *compositional* – compositional processes rely on a limited number of simple elements (e.g. simple concepts) that can be flexibly combined and recombined to produce more complex elements (e.g. complex concepts) according to certain rules. In other words, compositionality means that the meaning of complex conceptual structures is a function of the meaning of more basic elements and the way they are combined (Fodor 1975; Fodor and Pylyshyn 1988). A simple, concrete example of compositionality is the following case: the meaning of the sentence »Mary loves John« is a compositional function of the grammar and meaning of the units in the sentence, »Mary«, »John« and »loves«, the way in which these are combined into the sentence. Another complementary feature of higher cognition is *systematicity*, which denotes that there exist systematic relations between cognitive elements (Fodor and Pylyshyn 1988). While it is an important area of debate, I will not investigate systematicity further here due to space constraints and will assume that compositionality is systematic.

If human thought is compositional, then internal models that underlie thought processes are compositional as well. *Language of thought* (LOT) (Fodor 1975; Fodor and Pylyshyn 1988; Piantadosi 2021; Piantadosi et al. 2012, Piantadosi 2016) is an influential theory in cognitive science that successfully empirically and computationally investigates the nature of internal models that support compositionally structured thought processes. According to LOT, humans possess a symbolically structured internal model that operates analogous to language (Fodor 1975; Piantadosi 2021). The mind composes complex thoughts by reassembling parts of the internal model expressed in a symbolic language of thought (Fodor 1975; Fodor and Pylyshyn 1988; Piantadosi 2021; Piantadosi et al. 2016).

If the mind is compositional, what are its most basic elements? One way to think about this is in terms of primitives, that is, basic cognitive operations and composition laws that are available to a cognitive system before the process of compositional (hypothesis) formation begins (Piantadosi et al. 2012). In other words, primitives are the key cognitive operations that are available at the onset of learning and that allow for the acquisition of concepts and conceptual systems. For example, in the context of learning how to count, Piantadosi (2012) suggests that a plausible set of primitives that children possess at the start of this process consists of three primitive functions – *singletons*, *doubletons*, *tripletons* – which test whether a set contains one, two, or three elements. Other primitive operations in the context of counting could also be *next*, *previous* and *equal*, which provide functions for moving forwards and backwards on the count list and checking if two numbers are equal (Piantadosi et al. 2012). Some other, more general primitives are Boolean operators, such as *and*, *or*, *not* and *if ... then* (Piantadosi et al. 2012). Whilst some primitives are used across cognitive domains, some other primitives are more domain specific, such as in the case of counting.

To sum up, internal models that support higher cognition need to be compositional and contain a number of primitives that allow for the construction of more complex compounds. For the purposes of this paper, the crucial question is whether internal models in PP are compositional.

### Internal models in PP are not richly compositional

At the beginning, I discussed how generative models in PP consist of hypotheses that are arranged at different hierarchical layers. What exactly do these hypotheses encode? They encode the hidden causes of the sensory patterns in the co-occurring feature observations. More formally, hypotheses have a prior distribution and a likelihood distribution that describes how well each hypothesis captures the observed sensory pattern. Thus, differences between hypotheses pertain to the differences in the likelihoods of sensory patterns occurring under different hypotheses. Given such formalisation of hypotheses in PP, it is not immediately clear how they could account for the compositional character of thought. To use an enhanced version of an example referred to earlier, »Mary loves John because they share core values«, we understand the meaning of this sentence instantaneously, without having to rely on observing the co-occurrence of features or associations that code for this particular occurrence at some level. We understand this sentence because we understand the meaning of »love« and »values« in some abstract sense, long dissociated from learned associations between sensory observations. PP does not specify how the encoded co-occurring feature observations translate to primitive operations, what these might be in the first place, and how the primitives combine to result in new concepts, abstractions and laws. Thus, hypotheses as postulated by PP might not be sufficiently compositional and hence unable to account for logically structured higher thought.

Coming from a similar angle, Williams (2020) argues that hypotheses in PP

are only *weakly* compositional. As with any statistical model, weak compositionality is achieved by the possible assignments of probabilities to values and assignment of values to hypothesis variables where different assignments correspond to different facts about the world. Generative models construed this way have the expressive power of propositional logic (Russell 2010). To understand the limitations of propositional logic, it should be contrasted with a more expressive formal system, first order (i.e. predicate) logic (Haaparanta 2009), as follows (Williams 2020, p. 1768): »In contrast to factored representations, the ontology of first-order logic comprises not just facts but *objects* and *relations*, thereby representing the world as having *things* in it that are *related* to each other, not just variables with values‘ (Russell and Norvig 2010, p. 58). As such, first-order logic decomposes a propositional representation (e.g. Hume is a philosopher) into subjects (e.g. Hume), n-place predicates (e.g. is a philosopher), and quantifiers (e.g. there is an x such that x is Hume and x is a philosopher). The resulting gains in expressive power are famously enormous.«

Williams’ (2020) conclusion is that higher order thought is richly compositional and that the minimal formal system that can sufficiently capture this kind of rich compositionality is predicate logic. Generative models in PP, in contrast, can be described by propositional logic, which can only account for weak compositionality. Therefore, generative models in PP are currently not sufficiently rich to fully account for higher order thought.

### Could PP account for the compositionality of thought?

In the following, I discuss how compositionality might be addressed within PP, and in particular, what compositionality precursors might look like in infant generative models. In the cognitive science community, it is a relatively non-contentious assumption that generative models do not need to be learned from scratch (Friston et al. 2021). That is, evidence exists that infants are born with some built-in evolutionary acquired cognitive architecture and functionality; they perform basic statistical computations (Kirkham et al. 2002), engage in probabilistic reasoning (Denison et al. 2013), have some expectations about the physical and the social world (Spelke and Kinzler 2007), can reason about hidden causes (Koster-Hale and Saxe 2013) and engage in model updating and prediction error minimisation (Kayhan et al. 2019; Zhang et al. 2019).

Additionally, infants less than one year old (Lewkowicz et al. 2018; Werchan et al. 2015) are already able to learn hierarchical rules and can engage in the precursors of analogical thought (Gentner and Hoyos 2017), which might be interpreted as an indicator of infants possessing some compositional capacity from very early on. Based on these empirical findings, it is plausible to assume that the basic structure of the *compositional* hierarchical generative model is already in place in infants. What might cognitive architectures that support compositional higher thought look like?

*Initial state of the generative model*

I propose that initially, a simple distinction might exist between a lower level and a higher level of a generative model and gradually through development, the model will become populated with more and more complex hypotheses at different hierarchical levels. I suggest that the higher level of minimally compositional infant models contains some core primitive functions – within PP, these would correspond to simple hypotheses. Such hypotheses would equip infants with a general representational capacity and would need to be sufficiently rich to create meaningful, more complex hypotheses at lower levels of the hierarchy when combined. Whilst the higher levels of the model would specify available primitives, the lower levels would represent the compositions of these primitives into concrete hypotheses that could explain the causes underlying the incoming sensory sensations.

Similarly, coming from a general Bayesian tradition (though not PP), Perfors (2012) suggests that within (adult) generative models, a distinction exists between a so-called *latent* hypothesis space and an *explicit* hypothesis space. The former contains general primitive functions that can be combined into more complex hypotheses – as such, it represents the space of all possible hypotheses that can be represented. These simple hypotheses in the latent hypothesis space can be combined to create more complex hypotheses for predicting and explaining the incoming sensory evidence – such hypotheses belong to the explicit hypothesis space. The crucial distinction between the two kinds of hypothesis spaces or levels is in the *type* and *function* of the hypotheses they possess. Hypotheses in the latent hypothesis space represent basic and general building blocks for building concrete hypotheses that can then be actively manipulated by a cognitive system, for example, for predicting and evaluating sensory evidence in the explicit hypothesis space (Perfors 2012).

*Possible primitives in initial generative models*

Next, I suggest two kinds of primitives – individual primitives and systems of primitives – that might be present in the highest levels of the earliest generative models.

One set of primitives that infants could possess would be involved in constructing hypotheses for explaining a wide range of sensory observations across domains; thus, these primitives would be domain-general. An important condition that such primitives would need to fulfil is that they would allow for emerging compositionality that parallels that of humans. To my knowledge, developmental scientists are yet to map out a list of such domain-general primitive functions in infants although there are findings that show that eight-month-old infants can learn the same/different relation (Addyman and Mareschal 2010) as well as empirical evidence that five-month-olds can perform simple addition and subtraction (Simon et al. 1995). These findings indicate the type of primitives that could be found in infant generative models. Another set of expressively powerful primitives



that allow for rich compositionality was suggested by Piantadosi (2016) in the context of adult concept learning. He found that adult concept learning can be sufficiently explained only by a small set of primitives, such as »and«, »or«, »not« and »if and only if«. Note that these are just some ideas for the kind of primitives we might be looking for in the infant mind.

The other set of theoretical ideas that could inspire what the nature of primitive hypotheses might be comes from research on *core knowledge* (Spelke et al. 1992; Spelke and Kinzler 2007). This research pertains to early emerging or possibly innate expectations about the world, such as expectations about object behaviour, agents, number, space and time (Spelke et al. 1992; Spelke and Kinzler 2007). Core knowledge is general and abstract and applies to every member within a domain and therefore can be thought of as laws and principles that govern the behaviour of entities in specific domains. Therefore, core knowledge is a kind of a start-up software or a start-up library of primitive operations and programs (combinations of primitive operations) built in by evolution (Ullman and Tenenbaum 2020). An example of such a primitive program would be the principle of *solidity* – 2.5-month-old infants already expect that solid moving objects will not pass through another solid object (Renee Baillargeon 2004). Similarly, 2.5-month-olds and 3.5-month-olds understand the principle of *continuity* (Renée Baillargeon 2008) and *object permanence* (Renée Baillargeon and DeVos 1991), respectively. Infants expect that objects continue to exist when occluded and seem surprised if objects magically disappear.

For a side-by-side comparison of both ideas for primitives at the highest levels of the generative model (as presented in this section) and how they might relate to complex hypotheses at the lower levels (as presented in *The initial state of the generative model*), see Figure 1.

(a)

<b>COMPOSITIONAL BUILDING BLOCKS FOR HYPOTHESES</b>	
<b>Individual primitives</b>	<b>Programs - collections of primitives</b>
Domain general	Mostly domain specific
Single operations (e.g. and, or, if ... then, add)	Laws governing specific domains (e.g. principle of continuity, object permanence, principle of goal-directedness)
Abstract	

Higher levels of generative model



Primitives and programs are used for generating concrete hypotheses.

(b)

<b>HYPOTHESES</b>
Composed from primitive hypotheses and programs
Applied to concrete sensory observations
Actively used in prediction and explanation

Lower levels of generative model

Figure 1. Higher-level primitives and lower-level concrete hypotheses. (a) Two types of primitives might co-exist in the highest levels of infant generative models: individual, general primitives and programs, which are combinations of primitives and explain the laws underlying individual cognitive domains. (b) Concrete hypotheses are formed at the lower levels of infant generative models: these hypotheses are formed based on the higher-level primitives and are constructed in response to concrete observations.

*What learning mechanisms allow for the development and refinement of these internal models?*

Learning is the result of generative capacity of the mind

The first part of this paper discusses the centrality of compositionality to human thought (Fodor 1975; Fodor and Pylyshyn 1988; Piantadosi 2016). In fact, compositionality is usually presented together with another feature that is central to higher cognition – *productivity* or *generativity* (Fodor 1975; Fodor and Pylyshyn 1988). Whilst compositionality characterises the nature of internal models that can support higher thought processes, productivity is specifically important in the context of learning and development. Productivity can most succinctly be characterised as the »infinite use of finite means«, which means that a cognitive system equipped with a productive capacity is, in principle, capable of producing an infinite amount of thoughts and concepts. An example of a productive system is language; English contains a finite number of words and because there is no upper limit on the length of sentences, there is also no upper limit on the number of unique sentences that can be formed. The capacity for a speaker’s sentence construction is productive, which means that they are able to, in principle at least, form an infinite number of unique sentences (Katz 2009).

## Do internal models in PP have generative capacity?

Productive capacity necessitates that there exist some sort of learning mechanism that can generate new thought, that is, hypotheses. What might this mechanism be in the context of PP? As discussed at the beginning, the basic mode of learning in PP that pertains to model changes is Bayesian model updating, where the incoming sensory evidence is continuously changing the probability distribution over *existing* hypotheses. Thus, casting learning as Bayesian model updating is suitable when all the relevant hypotheses are already present in the model; however, it cannot explain how novel hypotheses are generated in the process of learning and development.

To address this exact problem, very recently another form of learning in PP has been formalised – *structure learning* (Friston et al. 2017; Kwisthout et al. 2017; Neacsu et al. 2022; Rutar et al. 2022; Rutar et al. 2023; Smith et al. 2020) – and empirically differentiated from Bayesian model updating (Rutar et al. 2023). This type of learning, unlike Bayesian model updating, allows for changing the structure of the generative model by removing, adding, merging hypotheses, or establishing new connections between hypotheses. Consider the following example from earlier (your friend always being fashionably late) that illustrates the difference between structure learning and Bayesian model updating: your friend continuously being late leads to gradual updating of the hypothesis (about your friend's arrival) such that after some time, the hypothesis starts predicting the friend coming at the proposed time with only a low probability. Then one day, your friend tells you that she is always late because she got two small puppies that always demand her attention when she is about to head out. What accounts for the acquisition of a new belief of why your friend is late is structure learning – your model added a new hypothesis »puppies attention« that is causally connected to the hypothesis coding for »friend being late«.

What is fascinating for the purpose of this paper is that structure learning could potentially offer a formal basis for addressing the productivity of thought from the PP perspective. If compositional models provide primitive elements, then structure learning could specify ways of combining these primitives to trigger new thoughts, concepts and ideas. More specifically, what I am suggesting here is that primitives combined in a new way would result in a person coming up with a concept that would correspond to a new hypothesis being added in a generative model via structure learning. Yet another example of structure learning could be to combine the primitives in a way that results in a more *detailed* hypothesis. Each new and meaningful combination of primitives would correspond to a distinct structural change in a model, such as adding a new hypothesis, removing an old one, merging existing ones, etc. Manipulating the hypothesis space via structure learning and hence manipulating the expressive power of the generative model could then be a PP way of satisfying the productivity condition.

The same mechanism – structure learning – could also explain developmental processes. For example, in the previous section, I suggested that initially, infants have a very crude structure of a generative model in place with the simplest

hypotheses – primitives – on the highest level of the model. Development with respect to structure learning could be characterised as a gradual population of the generative model with hypotheses at different levels. More specifically, one aspect of the emerging generative model might be important in the context of development: establishing hypotheses at different levels of *granularity*. Granularity, or the level of detail of hypotheses, pertains to the amount of information present in one hypothesis (Kwisthout et al. 2017). For example, a hypothesis with a high level of detail would be that when I go to a dog shelter, I will see a brown puppy Labrador, and a hypothesis with a low level of detail would be that I will see a Labrador. Learning different levels of granularity of hypotheses might underlie cognitive processes such as category learning in children, in which more abstract categories would correspond to hypotheses with a lower level of detail and learning sub-categories would pertain to more detailed hypotheses. As children’s generative models mature, they might start additionally considering the prediction error–information trade-off when choosing the appropriate level of detail of a hypothesis. The trade-off pertains to the fact that the more detailed the hypothesis, the more information it can yield if correct, but such a hypothesis can also generate more prediction error if wrong (Kwisthout and van Rooij 2015). For example, if I am not sure that the dog shelter allows for puppies, then predicting that I will see a brown puppy Labrador is costlier (as my prediction will be likely wrong and hence generate prediction error) than predicting that I will see a Labrador. Development thus pertains not only to generating increasingly more detailed hypotheses but also to choosing hypotheses with the appropriate level of detail.

## **Predictive processing in a pedagogical context**

In this final chapter, I sketch a reciprocal relationship between pedagogical practices and PP as a theoretical framework.

### *Pedagogical practices play a crucial role in model building*

The focal point of this paper has been to provide PP theoretical foundations for studying the development of higher cognition and to elucidate the learning processes involved in this development. What is at the core of learning and development is »change« or »transition« within one’s internal model. Learning can be characterised as moving from a place of not knowing to a place of knowing (more), and development is about a transition from a less developed to a more developed cognitive capability. Whilst in this paper I have, so far, focused exclusively on the theoretical characterisation of this change in an individual, crucially, learning and developmental changes most often happen in (in)formal learning environments. Thus, pedagogical practices, which create learning environments, play an important role in an individual’s model building. Below, I present three ways in which individual learning and teaching are intertwined and point out how internal model building might benefit from pedagogical guidance.

*Goal setting:* Learning is a directed activity in which a learner tries to get from an initial point of not knowing to a more advanced point of knowing. Often-times, this is achieved through explicit goal setting. Well-defined goals can specify the structure of the problem as well as the solution (Magid et al. 2015). Such structuring makes learning more manageable as it constrains a student's hypothesis space, that is, the space of all hypotheses that a student might consider in solving a problem (Magid et al. 2015). Guidance from teachers can help shape students' learning (by the above-mentioned constraining of the hypothesis space) and in so doing specify where in the learning space learning progress should happen. It is as if well-defined goals carve out the path for learning in a student's model.

*Scaffolding:* If goal setting defines the starting and end point of learning, another crucial ingredient is finding the right steps to get from the starting to the end point. Model building is a complex cognitive activity in which the structured sequence of steps matters a lot. Well-structured learning environments will have a greater potential to incur well-connected and meaningful structural changes in pupil's internal models (Vosniadou et al. 2001). Namely, when a model is in its immature state and hence a student does not possess a lot of knowledge in one domain yet, the student cannot just 'jump' to a highly advanced, complex model, but rather needs to take systematic, ordered steps. And this is exactly where teachers play a highly valuable role, as they can a) identify the starting level of pupils and b) scaffold their learning process, that is, challenge them just a bit above their current level of competence and help them get to the next competence level through guided instruction and probing questions (Vygotsky and Cole 1978). If goal-setting carves out the learning space, then continuous and gradual scaffolding guides children systematically through this space.

*Feedback:* Evidence about the role of explicit feedback in teaching is mixed and its efficacy can depend on how and when it is delivered and what is corrected (see Ellis 2009, for an overview). I suggest that in the context of model building, teacher's feedback might play a crucial role as it can reinforce successful cognitive strategies and promote meta-cognitive learning skills in students. Teachers can, in delivering their feedback, elaborate why certain ways of solving problems are better than others, and they can also encourage pupils to actively reflect on their own problem-solving approaches. In the previous section, I argued that teachers can *implicitly* scaffold learning for students via guided instruction, whereas feedback could be regarded as an *explicit* form of scaffolding.

### *Predictive processing as a theoretical tool for pedagogical practices*

Above, I show how pedagogical practices are constitutive of model building (in PP). Now, the reverse relationship also holds true. Namely, PP is a useful theoretical tool for understanding learning and teaching processes. Often, references to the »active student«, »active learning«, »student-oriented approach«, etc. are made only at the level of phraseology and blanket assertions. To substantiate these claims, it is important to elaborate on these concepts precisely and shed light on

the neural and cognitive mechanisms that underpin them. This is exactly where I see the contribution of a) the PP framework and specifically of b) my theoretical framework (for characterising the development of higher cognition through PP).

Regarding a), according to PP, humans are continuously probabilistically predicting their world (Clark 2013; Friston et al. 2016; Hohwy 2013), gathering new evidence so as to reduce prediction error and associated uncertainty (Clark 2013, 2015; Friston and Kiebel 2009; Hohwy 2013), and updating their beliefs according to new evidence (Oaksford and Chater 2009; O'Reilly et al. 2013). Recent experiments with cognitively inspired robots (which are in line with human experiments, e.g. Poli et al. 2020) have also found that the most effective algorithm for self-organisation and structuring of the learning curriculum in novel environments is the maximisation of »learning progress« (Oudeyer and Smith 2016). That is, when the robot was equipped with the algorithm that maximised learning progress, it avoided activities that were too hard to learn as well as those that were too easy to learn, instead exploring activities that yielded the biggest learning progress. When the learning progress of the current activity started to plateau, the robot switched to the activity that offered the second biggest learning progress. All these results suggest that humans are intrinsically highly active learners that are continuously adjusting their internal models, seeking out novel information and actively structuring their learning environments.

In one sense, any mental act – such as gathering information, updating and minimising prediction error – is an activity. However, a different kind of activity is also involved in learning, that is, higher cognitive processes. Such cognitive activities play a constitutive role in effectively executing and regulating the learning progress through the use of abstraction, generalisation, induction, deduction, classification, analogy, etc. (Ritchhart et al. 2011). Therefore, regarding b), my PP theoretical advance on the development of higher cognition is not only interesting from a theoretical point of view (expanding the explanatory scope of PP), but also provides the cognitive and possibly neural basis (as PP is neuroscientifically sound) of higher cognitive processes at work in learning. Whereas cognitive activities described under a) proceed without any instruction or guidance, higher cognitive activities described under b), crucially, rely on teachers' support in the form of scaffolded instructions, questioning and supportive guidance.

## Conclusion

PP promises to provide a unifying account of all human functioning, from perception and action to cognition. Whilst PP has been successfully applied to various aspects of human functioning (Friston and Penny 2011; Heil et al. 2019; Kaplan and Friston 2018; Rutar et al. 2023; Summerfield and de Lange 2014) its application to higher order cognition and development has been slower. Some researchers have shown, for example, that the basic PP machinery, that is, Bayesian model updating (Kayhan et al. 2019; Kayhan et al. 2019) and prediction error minimisation (Zhang et al. 2019), is already in place in infants. Likewise, Köster

(2020) has proposed ways in which PP is compatible with multiple developmental findings. In addition, researchers have recently started evaluating higher cognition from the PP perspective (Heil et al. 2019; Neacsu et al. 2022; Rutar et al. 2022; Rutar et al. 2023; Smith et al. 2020). In short, in recent years, the foundations have been laid for rigorously examining the empirical and modelling basis of the PP account of development and higher cognition. However, what we are still lacking is a mechanistic account, an overarching theory of higher-order cognition and a theory of development (though see Ward et al. under review, as a first attempt at constructing a theory of development in PP), or even more ambitiously, a theory of the development of higher cognition.

In this paper, I provided some initial elements that might be important to consider in developing a full-blown theory of the development of higher order thought in PP. I suggested that two questions should be considered in this context – What is the nature of internal models that support higher-order thought? And what learning mechanisms allow for the development and refinement of these internal models? – and proposed that internal models that support higher cognition are *compositional* and the learning mechanism that further refines and develops internal models needs to possess *generativity* (Fodor 1975; Fodor and Pylyshyn 1988). I further argued that PP could be complemented with LOT to account for the compositionality condition and that structure learning could be a good formalism to capture the generativity of thought.

Future work should investigate the suggested leads further and, most importantly, evaluate the compatibility between conceptual contributions here and the mathematical/formal basis of PP. Additionally, if within PP the core aim of cognitive systems is to minimise prediction error in the long run, the obvious question in the context of higher cognition is how higher thought processes such as thinking, imagining and learning new concepts fulfil that goal. Importantly, I have also argued that considering the role of education and teaching will be important in the context of understanding how infants' internal models develop. This last point calls for a tighter link between cognitive science, educational policies and teaching practice – how exactly findings from this research should inform educational practices should be investigated in the future. Many open questions remain; however, I believe that constructing a PP theory of the development of higher cognition will be worthwhile.

## References

- Addyman, C. and Mareschal, D. (2010). The perceptual origins of the abstract same/different concept in human infants. *Animal cognition*, 13, pp. 817–33.
- Baillargeon, R. (2004.) Infants' reasoning about hidden objects: Evidence for event-general and event-specific expectations. *Developmental science*, 7, issue 4, pp. 391–414.
- Baillargeon, R. (2008). Innate ideas revisited: For a principle of persistence in infants' physical reasoning. *Perspectives on psychological science*, 3, issue 1, pp. 2–13.
- Baillargeon, R. and DeVos, J. (1991). Object permanence in young infants: Further evidence. *Child development*, 62, issue 6, pp. 1227–46.

- Bayes, T. (1958). Studies in the history of probability and statistics: Ix. Thomas Bayes' essay towards solving a problem in the doctrine of chances. *Biometrika*, 45, issue 3, pp. 296–315.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and brain sciences*, 36, issue 3, pp. 181–204.
- Clark, A. (2015). *Surfing Uncertainty: Prediction, action, and the embodied mind*. Oxford: Oxford University Press.
- Denison, S., Reed, C. and Xu, F. (2013). The emergence of probabilistic reasoning in very young infants: Evidence from 4.5- and 6-Month-Olds. *Developmental psychology*, 49, issue 2, p. 243.
- Ellis, R. (2009). Corrective feedback and teacher development. *L2 Journal*, 1, issue 1.
- Fabry, R. E. (2017). Predictive processing and cognitive development. In: T. Metzinger and W. Wiese (Eds.). *Philosophy and predictive processing*: 13. Frankfurt am Main: MIND Group.
- Fodor, J. A. (1975). *The language of thought*. Vol. 5. New York: Harvard University Press.
- Fodor, J. A. and Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28, issue 1/2, pp. 3–71.
- Friston, K. J. (2009). The free-energy principle: A rough guide to the brain? *Trends in cognitive sciences*, 13, issue 7, pp. 293–301.
- Friston, K. J. (2010). The free-energy principle: A unified brain theory? *Nature reviews neuroscience*, 11, issue 2, pp. 127–38.
- Friston, K. J., FitzGerald, F., Rigoli, F., Schwartenbeck, P. and Pezzulo, G. (2016). Active inference and learning. *Neuroscience & biobehavioral reviews*, 68, pp. 862–79.
- Friston, K. J., Lin, M., Frith, C. D., Pezzulo, G., Hobson, J. A. and Ondobaka, S. (2017). Active inference, curiosity and insight. *Neural computation*, 29, issue 10, pp. 2633–83.
- Friston, K. J. and Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical transactions of the royal society B: Biological sciences*, 364, issue 1521, pp. 1211–21.
- Friston, K., J. Moran, R. J., Nagai, Y., Taniguchi, T., Gomi, H. and Tenenbaum, J. (2021). World model learning and inference. *Neural networks*, 144, pp. 573–590.
- Friston, K. J. and Penny, W. (2011). Post hoc Bayesian model selection. *Neuroimage*, 56, issue 4, pp. 2089–99.
- Gentner, D. and Hoyos, C. (2017). Analogy and abstraction. *Topics in cognitive science*, 9, issue 3, pp. 672–93.
- Haaparanta, L. (2009). *The development of modern logic*. New York: Oxford University Press.
- Heil, L., Colizoli, O., Hartstra, E., Kwisthout, J., van Pelt, S., van Rooij, I. and Bekkering, H. (2019). Processing of prediction errors in mentalizing areas. *Journal of cognitive neuroscience*, 31, issue 6, pp. 900–912.
- Hohwy, J. (2013). *The predictive mind*. Oxford: Oxford University Press.
- Kaplan, R. and Friston, K. J. (2018). Planning and navigation as active inference. *Biological cybernetics*, 112, issue 4, pp. 323–43.
- Katz, M. (2009). The language of thought hypothesis. *Internet encyclopedia of philosophy*, pp. 1–10.
- Kayhan, E., Heil, L., Kwisthout, J., van Rooij, I., Hunnius, S. and Bekkering, H. (2019). Young children integrate current observations, priors and agent information to predict others' actions. *PloS One*, 14, issue 5, p. e0200976.
- Kayhan, E., Meyer, M., O'Reilly, J. X., Hunnius, S. and Bekkering, H. (2019). Nine-month-old infants update their predictive models of a changing environment. *Developmental cognitive neuroscience*, 38, p. 100680.



- Kirkham, N. Z., Slemmer, J. A. and Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, 83, issue. 2, p. B35–42.
- Kok, P., Brouwer, G. J., van Gerven, M. A. J. and de Lange, F. P. (2013). Prior expectations bias sensory representations in visual cortex. *Journal of neuroscience*, 33, issue 41, pp. 16275–84.
- Köster, M., Kayhan, E., Langeloh, M. and Hoehl, S. (2020). Making sense of the world: Infant learning from a predictive processing perspective. *Perspectives on psychological science*, 15, issue 3, pp. 562–71.
- Koster-Hale, J. and Saxe, R. (2013). Theory of mind: A neural prediction problem. *Neuron*, 79, issue 5, pp. 836–48.
- Kwisthout, J., Bekkering, H. and 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, pp. 84–91.
- Kwisthout, J., and van Rooij, I. (2015). Free energy minimization and information gain: The devil is in the details. *Cognitive neuroscience*, 6, issue 4, pp. 216–18.
- Lewkowicz, D. J., Schmuckler, M. A. and Mangalindan, D. M. J. (2018). Learning of hierarchical serial patterns emerges in infancy. *Developmental psychobiology*, 60, issue 3, pp. 243–55.
- Magid, R. W., Sheskin, M. and Schulz, L. E. (2015). Imagination and the generation of new ideas. *Cognitive development*, 34, pp. 99–110.
- Muckli, L., De Martino, F., Vizioli, L., Petro, L. S., Smith, F. W., Ugurbil, K., Goebel, R. and Yacoub, E. (2015). Contextual feedback to superficial layers of V1. *Current biology*, 25, issue 20, pp. 2690–95.
- Nagai, Y. (2019). predictive learning: Its key role in early cognitive development. *Philosophical transactions of the royal society B*, 374, issue 1771, p. 20180030.
- Neacsu, V., Mirza, M. B., Adams, R. A. and Friston, K. J. (2022). Structure learning enhances concept formation in synthetic active inference agents. *Plos one*, 17, issue 11, p. e0277199.
- Oaksford, M. and Chater, N. (2009). Précis of bayesian rationality: The probabilistic approach to human reasoning. *Behavioral and brain sciences*, 32, issue 1, pp. 69–84.
- O'Reilly, J. X., Schüffelgen, U., Cuell, S. F., Behrens, T. E. J., Mars, R. B. and Rushworth, M. F. S. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the national academy of sciences*, 110, issue 38, pp. E3660–69.
- Oudeyer, P. and Smith, L. B. (2016). How evolution may work through curiosity of driven developmental process. *Topics in cognitive science*, 8, issue 2, pp. 492–502.
- Perfors, A. (2012). Bayesian models of cognition: What's built in after all? *Philosophy compass*, 7, issue 2, pp. 127–38.
- Piantadosi, S. T. (2021). The computational origin of representation. *Minds and machines*, 31, pp. 1–58.
- Piantadosi, S. T., Tenenbaum, J. B. and Goodman, N. D. (2012). Bootstrapping in a language of thought: A formal model of numerical concept learning. *Cognition*, 123, issue 2, pp. 199–217.
- Piantadosi, S. T. (2016). The logical primitives of thought: empirical foundations for compositional cognitive models. *Psychological review*, 123, issue 4, p. 392–424.
- Poli, F., Serino, G., Mars, R. B. and Hunnius, S. (2020). Infants tailor their attention to maximize learning. *Science advances*, 6, issue 39, p. eabb5053.
- Ritchhart, R., Church, M. and Morrison, K. (2011). *Making thinking visible: How to promote engagement, understanding, and independence for all learners*. San Francisco: John Wiley & Sons.

- Russell, S. J. (2010). *Artificial intelligence a modern approach*. New Jersey: Pearson Education, Inc.
- Rutar, D., Colizoli, O., Selen, L., Spieß, L., Kwisthout, J. and Hunnius, S. (2023). Differentiating between Bayesian parameter learning and structure learning based on behavioural and pupil measures. *PloS one*, 18, issue 2, p. e0270619.
- Rutar, D., de Wolff, E. van Rooij, I. and Kwisthout, J. (2022). Structure learning in predictive processing needs Revision. *Computational brain & behavior*, 5, issue 2, pp. 234–43.
- Simon, T. J., Hespos, S. J. and Rochat, P. (1995). Do infants understand simple arithmetic? A replication of wynn (1992). *Cognitive development*, 10, issue 2, pp. 253–69.
- Smith, R., Schwartenbeck, P., Parr, T. and Friston, K. J. (2020). An active inference approach to modeling structure learning: Concept learning as an example vase. *Frontiers in computational neuroscience*, 14, issue 41.
- Spelke, E. S., Breinlinger, K., Macomber, J. and Jacobson, K. (1992). Origins of knowledge. *Psychological review*, 99, issue 4, p. 605.
- Spelke, E. S. and Kinzler, K. D. (2007). Core knowledge. *Developmental science*, 10, issue 1, pp. 89–96.
- Srinivasan, M. V., Laughlin, S. B. and Dubs, A. 1982. Predictive coding: A fresh view of inhibition in the retina. *Proceedings of the royal society of London. Series B. Biological sciences*, 216, issue 1205, pp. 427–59.
- Summerfield, C. and de Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. *Nature reviews neuroscience*, 15, issue 11, pp. 745–56.
- Ullman, T. D. and Tenenbaum, J. B. (2020). Bayesian models of conceptual development: Learning as building models of the world. *Annual review of developmental psychology*, 2, pp. 533–58.
- Urgen, B. A. and Miller, L. E. (2015). Towards an empirically grounded predictive coding account of action understanding. *Journal of neuroscience*, 35, issue 12, pp. 4789–91.
- Vosniadou, S., Ioannides, C., Dimitrakopoulou, A. and Papademetriou, E. (2001). Designing learning environments to promote conceptual change in science. *Learning and instruction*, 11, issue 4/5, pp. 381–419.
- Vygotsky, L. S. and Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Cambridge, MA: Harvard University press.
- Werchan, D. M., Collins, A. G. E., Frank, M. J. and Amso, A. (2015). 8-Month-old infants spontaneously learn and generalize hierarchical Rules. *Psychological science*, 26, issue 6, pp. 805–15.
- Williams, D. (2020). Predictive coding and thought. *Synthese*, 197, issue 4, pp. 1749–75.
- Zhang, F., Jaffe-Dax, S. Wilson, R. C. and Emberson, L. E. (2019). Prediction in infants and adults: A pupillometry study. *Developmental science*, 22, issue 4, p. e12780.

Danaja RUTAR (Nizozemska)

## RAZVIJANJE VIŠJIH KOGNITIVNIH PROCESOV SKOZI NAPOVEDNO PROCESIRANJE

**Povzetek:** Napovedno procesiranje je vplivna teoretična paradigma v kognitivni znanosti, ki obljublja poenotenje vseh vidikov človeškega delovanja: percepcije, motoričnega delovanja in kognicije. Teorija je sicer že uspešno razložila številne vidike zaznavanja in motoričnega delovanja, a ji, ključno, manjka koherentna teoretična razlaga kognitivnega razvoja in višjih kognitivnih procesov ter še bolj ambiciozno, razlaga *razvoja višjih kognitivnih procesov*. V tem prispevku raziskujem, v kolikšni meri je teorija prediktivnega procesiranja v sedanji obliki sploh primerna za razlago razvoja in višjih kognitivnih procesov. Predlagam, da glede na sedanje stanje napovedno procesiranje ne more dovolj razložiti razvoja višjih kognitivnih procesov, saj mentalnim modelom, ki jih predlaga, manjkata dve ključni sestavini, ki sta sicer pomembni za karakterizacijo višjih kognitivnih procesov: kompozicionalnost in generativnost. V tem članku raziskujem, kako bi lahko ti dve ključni funkciji obravnavali z vidika teorije prediktivnega procesiranja. Predlagam tudi, kakšna je vloga izobraževalnih in pedagoških praks v tem kontekstu. Nazadnje pa tudi razložim vzajemno razmerje med pedagoškimi praksami in napovednim procesiranjem: pedagoške prakse so ključnega pomena za izgradnjo mentalnih modelov, napovedno procesiranje pa nudi teoretični okvir za razumevanje psiholoških mehanizmov prisotnih v pedagoških kontekstih.

**Ključne besede:** napovedno procesiranje, razvoj višjih kognitivnih procesov, kompozicionalnost, generativnost, pedagoške implikacije

**Elektronski naslov:** danaja.rutar@gmail.com