



Review



Cite this article: Versace E, de Bivort BL. 2026 Multiple weak biases support adaptive choices without prior experience: a self-supervised strategy. *Proc. R. Soc. B* **293**: 20251878. <https://doi.org/10.1098/rspb.2025.1878>

Received: 23 July 2025

Accepted: 11 December 2025

Subject Category:

Behaviour

Subject Areas:

behaviour, cognition, neuroscience

Keywords:

predispositions, inductive bias, precocial, self-supervised strategy, unsupervised strategy, newborn, neonate, fixed-action pattern

Author for correspondence:

Elisabetta Versace

e-mail: e.versace@qmul.ac.uk

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.8230302>.

Multiple weak biases support adaptive choices without prior experience: a self-supervised strategy

Elisabetta Versace¹ and Benjamin L. de Bivort²

¹School of Biological and Behavioural Sciences, Centre for Brain and Behaviour, Queen Mary University of London, London, UK

²Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

EV, 0000-0003-4578-1851; BLdB, 0000-0001-6165-7696

From social interaction to foraging, the survival of newborn animals depends on their ability to make adaptive decisions without prior experience. While ethological research focused on strong innate biases, recent empirical evidence shows that early preferences are often weak, transient and variable. After reviewing these findings, we introduce a novel model that explains how multiple weak biases can support adaptive choices in naive, inexperienced animals. We propose a self-supervised strategy in which naive animals combine internal predispositions with environmental structure to guide behaviour. First, predispositions target features that are rare in the background but common in relevant stimuli (e.g. red colour, upward movement), thus reducing false positives. Second, the evolution of multiple weak biases allows animals to increase accuracy of decisions using co-occurring cues (e.g. colour, sound and movement). Weak biases are advantageous as they prevent strong responses to misleading partial matches. Requiring the co-occurrence of multiple rare cues supports effective choices in complex environments with many misleading stimuli. This simple self-supervised strategy offers a new framework for understanding early decision-making, with implications that span across behavioural ecology, developmental psychology, comparative cognition and artificial intelligence, where decision-making with sparse evidence is a crucial challenge.

1. Advantages and challenges of early-life biases

From approaching social partners [1,2] to predator avoidance [3] and foraging [4,5], the fitness of young animals relies on adaptive choices. While later-life decisions are influenced by learning, culture and plasticity [6,7], early choices occur independently of these factors. Inexperienced individuals can use biases evolved in previous generations to guide a range of preferences and decisions, before having learning opportunities (§2). However, these biases may increase the likelihood of errors [8], such as overlooking or misjudging critical information [9]. Errors due to biases are frequent especially when environments incur rapid changes [7,10], like in the case of sea turtle hatchlings disoriented by recently introduced artificial lights [11]. Balancing benefits and costs of biases is particularly important for young animals. *Precocial* animals [12] —species that move autonomously within hours after hatching or birth, including fowls, turtle hatchlings and many ungulates [13] —and solitary animals such as most insects [4] are exposed to these risks. See [table 1](#) for a glossary of key words in *italics*. We focus on precocial animals to understand how inexperienced individuals are cognitively and perceptually predisposed to leverage biases while deploying strategies to mitigate associated costs. At the same time, this reflection is relevant to understand the

role of *inductive biases*—a concept connected to that of innate knowledge—and built-in structure in artificial intelligence [14–16], as it is still unclear how to best tailor the architecture of artificial intelligence to enhance ‘adaptive’ choices in the absence of previous experience.

Evidence from different vertebrate species—including domestic chicks, tortoise hatchlings and human infants—show how early behavioural responses are guided by approach/avoid biases known as *early predispositions* [2,4,17,18]. In §3, we examine evidence that early predispositions are weak, flexible and transient, unlike the rigid, species-specific fixed-action patterns classically described by ethologists [19,20] and unlike the innate escape responses that require extensive evidence to be modified [21]. The widespread occurrence of multiple weak biases/early predispositions across species suggests that they serve an adaptive function that remains to be explored.

We hypothesize that the presence of multiple weak biases is an adaptive strategy for self-supervised decision-making. This strategy enables inexperienced animals to reduce errors in their biased choices by requiring the simultaneous presence of multiple corroborating cues before generating reliable behaviours or consolidated memories. In contrast, strong biases towards single cues, which act as triggers or releasers, are particularly useful in time-critical contexts, such as mating [9,19], fight-or-flight reactions [3,21,22] or orientation in stable environments [11]. We formalize this idea with a mathematical model of cue integration, evaluating how weak biases can serve as an adaptive solution for inexperienced organisms. Our model provides a set of testable predictions.

2. A multitude of biases in newborn animals

The spontaneous preferences observed in newborn animals with minimal or no prior experience offer a window into the preparedness of the nervous system to respond to external stimuli. Growing evidence not only challenges the old empiricist (Hume, Locke) and behaviourist (Watson, Skinner) claims that internal worlds are solely shaped by experience but also questions more recent accounts, mostly related to artificial intelligence development, that point at general-purpose learning mechanisms [14,23–26]. Far from being blank slates, young animals exhibit a multitude of biases [2,4,17,18,27,28] and prepared responses [29,30]. Similarly, even artificial networks, for which supervised learning is essential to their success, are pre-tuned to specific information via their structure and hyperparameters. For instance, in convolutional neural networks the model assumes that nearby pixels are related (with an induction for spatial locality) and that objects appear across space [31]. In recurrent neural networks the model assumes that the past influences the present (with a temporal inductive bias) [32]. See [14] for a review of inductive biases in artificial neural networks. Biases available to inexperienced systems do not need to be encoded as prescriptive rules or isomorphic representations, but can equally work as simple enhancers of information flow. Whether young animals use specialized mechanisms to refine the use of biases and reduce related errors in the absence of experience remains an open question.

As precocial [12,17] animals that benefit from the ability to accurately perceive stimuli and move autonomously within hours of hatching,¹ newly hatched chicks (*Gallus gallus*) have become a valuable system to study early predispositions, the early biases that support approach and withdrawal responses in newborn animals, and inductive biases [18,24,36]. Unlike mammals with internal gestation, chick embryos can be incubated under controlled conditions, allowing researchers to strictly regulate their sensory experiences in the laboratory. Moreover, chicks can be hatched and raised individually, or with the social support of stimuli controlled by the experimenters, and tested while still inexperienced, or while having been exposed only to specific stimuli until the moment of test. Altricial species, on the contrary, require more extensive parental care. These characteristics have enabled researchers to identify and study numerous predispositions in poultry chicks (for recent overviews see [2,17,18,27,37]).

As evident in poultry chicks, early biases can facilitate rapid learning via *imprinting* [38–41]. Imprinting is a fast-learning mechanism that depends solely on exposure, by which chicks learn the distinctive characteristics of the first stimuli they are exposed to and develop attachment behaviours for them. In the wild, the first encountered stimuli are typically the mother and siblings [38–40,42]. Notably, predisposed (spontaneous) and imprinted (learnt) preferences rely at least partially on different neural substrates, as shown by electrophysiological [43], lesion [44], pharmacological and immunohistochemical studies [45,46]. This reinforces the idea that early biases serve as an orientation mechanism for inexperienced individuals, before learning takes place [18]. Neuroscientists are discovering the neuroanatomical and cellular mechanisms that underlie the interactions between innate and learned responses, for instance in escape behaviour [21]. Before animals are engaged in learning, predispositions are essential to orienting young animals towards ‘appropriate’ objects, ensuring that learning and attachment are directed at objects that enhance survival and fitness [2,17,18,47]. Interactions between biases and learning are relevant in several domains and throughout life, as shown by efforts in biodiversity conservation projects [10]. In this article, we focus on early choices of naive individuals and model the example of newborn chicks that choose which stimuli to approach.

Studies on visually naive chicks have revealed several population-level preferences for both static and dynamic visual features. These include preferences for static patterns such as face-like stimuli [43,47–49], hollow objects [50], specific colours (notably red and blue, which are rare in the background landscape [36,51,52]), a stuffed hen versus an unstructured reassembly of its components [28,49,53], intermediate size stimuli that match the size of a pre-domestication fowl [54], with stronger

¹In many songbirds, predispositions for species-specific songs guide learning [33]. Studying these predispositions is methodologically challenging, as it requires to isolate young birds and to prevent them hearing their own vocalizations. Because these *altricial* birds hatch with immature sensory, perceptual and motor systems, test in free-ranging tasks without caregiver interference is often unfeasible. Similar constraints apply to human infants. While research suggests that language and music acquisition in humans is influenced by innate biases [34,35], these studies are typically conducted years after birth, complicating inferences about early development.

Table 1. Glossary

<i>Additive models</i> : models in which a decision is determined by the summed contributions of the preference for individual cues, without interactions or scaling effects between them
<i>Altricial</i> : species in which newborn animals have sensory and motor systems not fully developed, with consequent limitations in sensation and movement
<i>Bias</i> : systematic tendency towards or against a stimulus or pattern. In our model it is represented by β
<i>Early predisposition</i> : bias that produces approach/withdrawal or preferential attention responses in newborn animals
<i>Fixed-action pattern</i> : non-learned stereotypical behavioural sequence triggered by specific stimuli, that usually runs to completion regardless of external changes
<i>Imprinting</i> : learning mechanism based on exposure during early time windows of development, that produces attachment for a salient object or location
<i>Inductive biases</i> : preferences in favour of some solutions over others in front of novel scenarios
<i>Precocial</i> : species in which newborn animals have highly functional sensory and motor systems, with advanced sensory and motor skills
<i>Self-supervised strategies</i> : cognitive/computational and behavioural strategies that rely on the intrinsic structure of the input, acting as implicit supervision
<i>Spontaneous biases</i> : systematic tendencies towards or against a stimulus, that do not depend on prior experience
<i>Synergistic models</i> : models in which a decision is determined by the interaction of preferences for different cues, including scaling and nonlinear effects between them. In our model, the scaling factor is represented by λ
<i>Unsupervised strategies</i> : cognitive/computational and behavioural strategies that rely on internal guidance (rather than on external instructions and reinforcement) to find solutions

preferences for certain *combinations* of colour and size [51]. Interestingly, spontaneous preferences are often stronger for dynamic than for static stimuli features: chicks prefer objects that move upward against gravity [37], biological motion [55], objects that move along the longer axis [56] or self-propel [57]. It has been argued [2,17,27,47] that many of these features are associated with the presence of social partners, or more in general the presence of living animals and *animacy* (the property of being alive) [1,27,58]. As vulnerable hatchlings rely on maternal care for protection and sustenance, the ability to orient towards stimuli appropriate for subsequent imprinting provides a clear adaptive advantage [18]. Abundance of information associated with living animals can also explain why even solitary species, such as land tortoises [1], exhibit similar biases observed in chicks [43] and infants [59,60]. *Altricial* animals, like newborn infants, have a very different course of development from precocial animals, as they are highly dependent on parental care and born with immature sensory and motor systems. Despite these differences, both exhibit remarkably similar biases, beyond face-like pattern sensitivity, such as the preferences for biological motion [55,61], self-propulsion [58], movement along the main body axis [62], as previously highlighted [2,17,18,27]. The translational value of poultry chicks as a model for early predispositions has also been previously discussed [63].

Overall, early biases are widespread among newborns of various species. This observation may initially suggest that such biases are robust, consistent and universally fixed to enhance adaptive success. However, as §3 demonstrates, these predispositions are neither rigid nor universal but rather weak and transient.

3. Biases with different strength: fixed-action patterns versus early predispositions

Biases that do not depend on learning can differ in strength, flexibility and consistency across a lifetime [28]. Biases can be characterized as coming from a continuum: high flexibility is often associated with weak biases, while lower flexibility is associated with strong biases (see figure 1), along a continuum. Biases with low flexibility and high strength have been described as *fixed-action patterns*, i.e. rigid behaviours that are intense, persistent and resistant to extinction, triggered by specific stimuli and often leading to striking errors such as attacking unrealistic dummies [20], incubating oversized and unrealistic eggs [20] or courting artefacts instead of other animals [9]. An organism that exhibits such errors may nevertheless have high fitness, if false negatives (failing to respond to a critical stimulus, missing reproduction due to a sexual competitor or lack of egg incubation) are more costly than false positives (responding to a non-target stimulus) [64]. Importantly, more recent evidence has shown some flexibility of fixed-action patterns in response to developmental experience and context (e.g. [65] for the aggressive responses of three-spine sticklebacks [66], and [67] for the flexibility in escape behaviour in mice and flies). Nevertheless, these behaviours are overall stable and require repeated evidence to be modified. An example of fixed-action pattern encountered in the first stages of life is the hatching behaviour of birds, which is highly conserved across species and responsive to proprioceptive and hormonal cues [68]. Other strong biases that hold at the population level and orient animal behaviour in different domains include phototaxis [11], and see [69] for a clear population bias in flies, sensitivity to cues that elicit predator avoidance [3,22,70], foraging [4,71,72], prey-catching [71] and oviposition [73,74].

At the other end of the continuum, we find spontaneous biased behaviours that are blurred and flexible. *Early predispositions*, a term used in both human and non-human animal cognition studies [2,18,44], refers to the consistent population preferences in orienting and approach responses [2,17,18,27] at the beginning of life. Interestingly, while these predispositions are observed in different experiments [28,49,53,75,76], they are by no means uniform and permanent. For instance, while most chicks are attracted by upward movement, biological motion, face-like patterns and hollow objects, other animals approach the alternative stimuli [37], explore both stimuli or do not approach any stimulus [37,49,50,55]. Moreover, the preference to approach an object that changes in speed (as if self-propelled) is consistently present in the first day after hatching, but is not observed 2 days later [77,78] (unless pharmacological treatment with T₃ hormones re-instantiate the original preference [78]), an example of

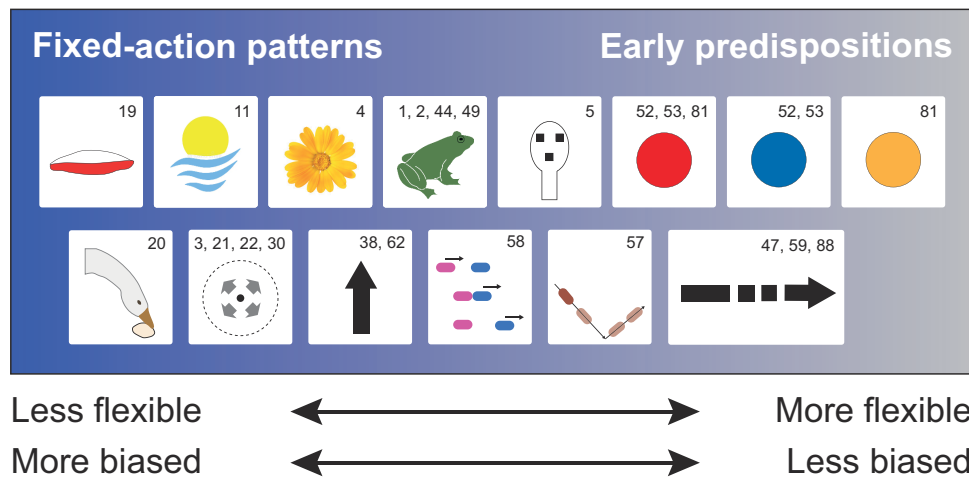


Figure 1. Innate biases vary in flexibility and strength. Fixed-action patterns usually have low flexibility and high strength, while early predispositions are weak and plastic. The figure illustrates some innate biases described in the references indicated with specific numbers. Cues position along this spectrum should be interpreted qualitatively. The exact position on the horizontal axis is purely indicative, as predispositions depend on species and context, as shown in this article.

temporal plasticity. Similarly, the initial preference to approach female conspecifics observed in the first hours after hatching, fades away [79]. In some cases, the initial predisposed preferences can be easily reversed by a simple prolonged exposure to the less preferred stimuli: in poultry chicks, the initial preference for approaching a red stimulus versus a yellow one can be reversed after prolonged exposure [36,80]. Additionally, predispositions are influenced by the context, as they are modulated and strengthened by the presence of co-occurrent stimulation [55,76,81,82], and variations in behavioural phenotypes can arise between individual animals and between breeds [53], rather than being a species-defining trait.

Unlike fixed-action patterns, which typically provide a specialized function for a specific circumstance, such as hatching and predatory threats, early predispositions function as flexible guiding mechanisms that shape learning and adaptation during early life. The weakness of early predispositions suggests that they act in a probabilistic manner, requiring multiple confirming cues to elicit strong responses. Young animals may integrate multiple weak biases as a *self-supervised strategy* to minimize errors. Rather than responding automatically to single stimuli, animals react when multiple cues co-occur, allowing them to exploit statistical regularities even without prior experience.

The adaptive value of predispositions is supported by different lines of evidence. First, predispositions are widespread across taxa, including birds, reptiles and mammals, and have been observed across different laboratories and settings, as exemplified by colour preferences [36,51,82,83]. Moreover, although predispositions are weak, they are conserved between breeds [53,77,78,84]. The soft, weaker nature of early predispositions compared with fixed-action patterns suggests that predispositions might guide behaviour requiring multiple confirming cues to elicit strong responses, as a safety net to avoid too-impulsive reactions. This aligns with the idea that young animals integrate multiple weak biases as a self-supervised strategy to minimize errors without requiring experience.

To summarize, biases range along a continuum from fixed-action patterns to early predispositions, serving as internal priors that shape responses to stimuli. Early predispositions are present shortly after birth in naive animals and are weak, transient and more flexible, allowing context-sensitive learning. Fixed-action patterns, in contrast, are more rigid, species-wide and less influenced by the environment, appearing in both young and experienced animals. These differences suggest distinct evolutionary roles: while fixed-action patterns ensure consistent responses to predictable stimuli, early predispositions guide learning towards variable but important stimuli.

4. Self-supervised sources of evidence

We investigate how multiple innate biases can be used before an animal engages in experience-dependent learning (e.g. individuals that have no previous experience with the sensory stimuli presented at the time of the behavioural decision) to reduce the risks of maladaptive responses, thus reducing the potential costs of those biases.

A plausible evolutionary strategy is the development of biases towards features that are statistically rare in the environment yet consistently present in biologically relevant targets. Such a bias would have a low false positive rate, and indeed several species use rare and conspicuous colours to communicate with conspecifics [85] and as sexual and warning signals [86]. Even in animals that have more cryptic, earth-toned plumage, such as female jungle fowls, the head region contains some reddish colours, rare in the general environment, that are attractive for juveniles [36,51]. Similarly, cues of animacy such as upward movement against gravity [37] or biological motion [55] are rare in inert objects. Once biases for rare features have evolved, they can be used synergistically to assess an overall scene composed of multiple cues or stimuli.

As a self-supervised strategy, co-occurrence of multiple favoured cues provides a more precise and reliable signal of stimulus value. Rather than assessing signal reliability by integrating information over time (e.g. continuously tracking a single cue over time or waiting for it to recur multiple times), animals can assess stimuli based on multiple instantaneous factors that enhance

decision-making in the single moment of the first choice. These factors can fall into a variety of categories including (i) features that are *rare in the background*, (ii) the use of context information, specifically the co-occurrence of multiple distinct cues and (iii) the level of independence of co-occurring cues (where co-occurring independent cues provide stronger information than co-occurring dependent cues). Requiring multiple corroborating cues before committing to a behavioural response allows for the presence of *weak* biases that individually would be too weak to produce consistent behaviours.

The presence of multiple independent features within a single stimulus is a condition that can be used to infer a higher degree of certainty in the validity of the stimulus. Independent features are often processed in different sensory pathways, such as colour and movement, or shape/size and colour. If two or more preferred characteristics co-occur despite being encoded independently (such as a ‘face-like’ [48,49] or ‘hen-like’ [28,53] configuration combined with self-generated movement [87], or an object displaying both a distinctive colour and biological motion [81] or an object that has multiple ‘adult-like’ features such as reddish colour and large size [51]), this convergence provides strong evidence that the stimulus is relevant, reducing the risk of errors in spontaneous approach behaviours. By prioritizing stimuli that simultaneously exhibit multiple, independently preferred traits, newborn animals can optimize early decision-making, ensuring that their first interactions with the environment are both efficient and adaptive.

5. Spontaneous choices model

(a) A model for spontaneous preferences in inexperienced animals

To investigate the role of multiple weak biases in newborn animals, we propose a mathematical hypothesis for how individuals with *spontaneous biases* but no prior experience can integrate multiple cues in decision-making (see electronic supplementary material, S1). As framing scenarios, we consider a newborn chick (or newborn insect) that faces the choice between one target stimulus such as the mother hen (or a target flower) versus alternative non-target stimuli, when multiple cues are present.

We focus on biases for cues that have low base probabilities (cues that are rare in a generic background), as other biases would be too costly evolutionarily, inducing too many false positives. The model tests whether relying on *multiple weak biases*, rather than a few strong ones, improves the accuracy of object selection. Specifically, we analyse the role of *strength* and *number of biases*, as well as the interaction of *valid cues* (preferred cues that correctly signal the presence of a target object) and *misleading cues* (preferred cues that are present in non-target objects). We compare choice scenarios in which cues contribute independently from one another (*additive model*) with those where they interact synergistically, with multiplicative and scaling effects (*synergistic model*). This provides a framework for understanding early predispositions as a self-supervised strategy that minimizes errors and maximizes choices’ adaptive value.

(i) Additive model

In the simplest scenario (the additive model), multiple cues contribute independently to decision-making, with each cue having a linear effect on the probability of selecting the correct object, based on cues that signal the presence of the target (valid cues) versus misleading cues that correspond to non-target objects. However, adaptive cues that have a low base probability (i.e. cues that occur less frequently in nature, such as blue or red colours, or upward movement against gravity) have a greater predicting value. Cues relevant for predisposed preferences include low-level features and configurations that are stable across individuals and contexts, retaining statistical regularity over evolutionary time. By contrast, individual and context-specific features require within-generation learning, and are not relevant for a model on predisposed behaviours. This model is given by the following equation:

$$P_{\text{additive}} = \frac{\sum_i [\beta_i (-\log_2(p_{c_i})) V_i]}{\sum_i [\beta_i (-\log_2(p_{c_i})) V_i] + \sum_j [\beta_j (-\log_2(p_{c_j})) M_j] + 1} \quad (5.1)$$

In this formulation, β_i represents the biases for the i valid cues that correctly signal the target object (e.g. a face-like pattern associated with the presence of the mother hen), and β_j is the biases for the j misleading cues. The use of indices i and j allows each cue to have distinct bias values and base probabilities, capturing potential differences in their individual strength or informativeness. Values of β greater than 1 signal bias for the stimuli, and values below 1 indicate repulsion. V represents the number of valid cues of type i , M represents the number of misleading cues of type j (e.g. a face-like pattern that occurs by chance or flipped face-like pattern); p_{c_i} and p_{c_j} are the base probabilities of valid and misleading cues, respectively, where less probable events carry a higher information value, as conveyed by the $-\log_2$ transformation [88]. The additional +1 term in the denominator accounts for baseline uncertainty.

In this model, the presence of multiple biases enables an individual to have weak biases to guide behaviour, reducing the costs of strong biases. However, each cue contributes linearly to the decision probability, meaning that the context of other cues present/absent does not enhance or diminish the effect of each cue. The ranges of parameters are presented in table 2.

(ii) Synergistic model

In many ecological settings, including the social attraction of newborn chicks for living beings and the first foraging choices of solitary animals, stimuli vary along multiple dimensions, such as colour, size, direction of movement for visual stimuli or pitch,

Table 2. Overview of parameters and range values for each parameter.

parameter	range
bias strength β	$(-\infty, \infty)$
synergy scaling factor λ	$\lambda \in [1, \infty)$
number misleading cues (associated with a non-adaptive target) M	$[0, \infty)$
base probability p	$[0, 1]$
number valid cues (associated with an adaptive target) V	$[1, \infty)$

timbre and duration for acoustic stimuli, or odour for olfactory stimuli. The *context* of surrounding data in which a cue appears can provide synergistic information on the reliability of the overall stimulus as a valid, adaptive target.

To model the synergy between cues that are simultaneously present, we introduce a *scaling factor* λ , which captures the degree of interaction between cues. When $\lambda > 1$, the contribution of cues is amplified, reflecting a synergistic effect. As in the additive case—where $\lambda = 1$ and no synergy is assumed—the influence of each cue is still modulated by any bias towards or against it, as well as its base probability. The success probability of a choice in this scenario is given by the following equation:

$$P_{\text{synergy}} = \frac{\sum_i [\beta_i (-\log_2(p_{c_i}) \cdot V_i)^{\lambda_i}]}{\sum_i [\beta_i (-\log_2(p_{c_i}) \cdot V_i)^{\lambda_i}] + \sum_j [\beta_j (-\log_2(p_{c_j}) \cdot M_j)^{\lambda_j}] + 1} \quad (5.2)$$

In this formulation, the λ_i term represents the synergy scaling factor for the co-occurrence of multiple valid or misleading cues. This model emphasizes the use of co-occurring cues as a self-supervised strategy, leveraging multiple biases and the correlation structure of the environment. In this framework, the presence of multiple cues that are correlated offers weaker synergistic evidence for a stimulus and would be assigned a modest value of λ (e.g. between 1 and 1.5). In contrast, convergence of cues from independent sensory modalities provides stronger evidence of stimulus validity as signal for a target, as the probability of such convergence occurring by chance is lower. Such multimodal interactions would therefore be associated with higher values of λ (e.g. 1.5–2), reflecting a stronger synergistic effect.

(iii) Model predictions

The model yields distinct and testable predictions on how naive, inexperienced animals integrate multiple cues when making spontaneous choices for a target stimulus. These predictions can guide experimental validation and parameter estimation.

In the additive model ($\lambda = 1$), cues contribute independently and linearly to the decision accuracy. The probability of a correct choice increases proportionally with the number of valid cues (V) and decreases with the number of misleading cues (M). Performance (choice of a valid, adaptive target) gains from adding more valid cues are constant. In this model, multimodal combinations or correlated cues do not enhance performance beyond the sum of their individual contributions. Therefore, behavioural data are expected to show steady, monotonic improvement in discrimination with increasing valid cues, but rapid decline when misleading cues accumulate. The model also predicts greater uncertainty and a lower probability of an adaptive response in conflicting situations, where valid and misleading cues are simultaneously present. A potential strategy that could be employed in such situations is deferring the decision, making time for more cues to occur. Additive integration is expected to prevail in contexts with low environmental noise.

In the synergistic model ($\lambda > 1$), cues contribute with supra-linear gains to the decision accuracy, particularly when valid cues co-occur across independent sensory modalities. The performance rises faster than linearly with the number of valid cues when these cues are correlated or multimodal. Misleading cues have a diminished impact, as synergy amplifies consistent evidence and suppresses conflicting signals. Therefore, synergistic, super-additive effects are expected, with performance exceeding the sum of single-cue effects. Strong synergy ($\lambda > 1.5$) is expected to maintain high accuracy even in the presence of several misleading cues, provided valid cues originate from different sensory modalities (e.g. visual and auditory). Therefore, the model predicts that spontaneous preferences will be stronger and more robust when stimuli present combinations of features that rarely co-occur in non-adaptive, misleading targets.

Overall, these predictions provide a quantitative framework for testing how multiple weak innate biases combine to guide early decision-making. Empirical data from controlled behavioural experiments or cross-modal stimulus manipulations can be used to estimate the parameters (β , λ , V , M , p) (summarized in [table 2](#)) and evaluate whether observed choices are better explained by additive or synergistic integration.

(b) Model behaviour

The model's behaviour under varying parameters is presented in [figure 2](#). The horizontal axis of each panel represents the number of valid cues (V) associated with the presence of an adaptive target, while different line colours indicate different numbers of misleading cues (M). Each panel shows a combination of synergy (λ) and bias strength (β). Synergy (λ) progressively increases in columns of subpanels from left to right, from the additive case ($\lambda = 1$) to stronger synergistic effects ($\lambda = 2$). The bias strength increases across rows of subpanels from no bias (β_i and $\beta_j = 1$) to strong bias (β_i and $\beta_j = 1$).

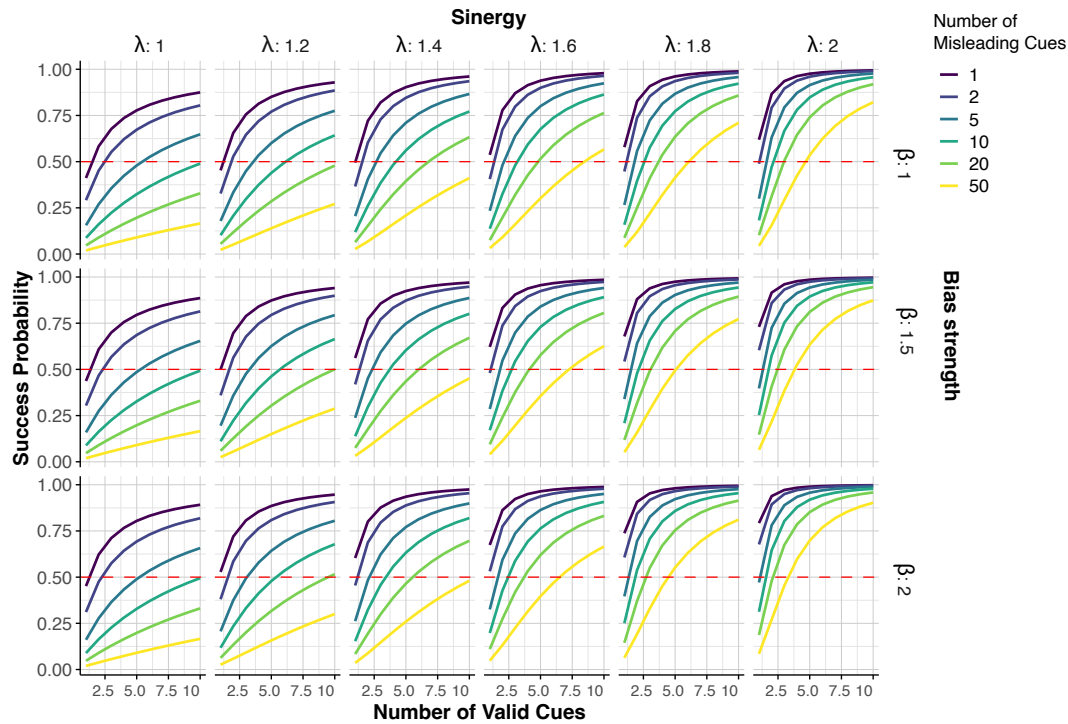


Figure 2. Success probability as a function of the number of valid cues, misleading cues, bias strength and cue synergy. The horizontal axis represents the number of valid cues, different line colours represent different numbers of misleading cues. Subpanel columns represent increasing levels of synergy (from $\lambda = 1$ for the additive model with no synergy, to synergy of 2). Subpanel rows indicate bias strength β .

In an *additive* scenario ($\lambda = 1$), cues contribute linearly to the decision process. Their influence is determined solely by their base probability and individual bias strength (β_i and β_j). In this regime, organisms perform poorly in conditions with many misleading cues. But additive integration is computationally (and potentially metabolically) efficient and robust in low-noise contexts, making it particularly suited to organisms with limited processing capacity or minimal environmental complexity. In newborn animals or AI systems with simple decision-making mechanisms, an additive strategy is computationally less demanding than tracking complex interactions between cues. However, as the number of misleading cues increases (indicated by lighter-coloured lines in figure 2), the linear model's success rate declines markedly due to the absence of a compensatory mechanism for noise suppression or feature reinforcement. This decline reflects an increase in uncertainty, which can translate not only into a lower probability of a correct choice but might also lead to longer hesitation or delayed decision-making, as the individual might attempt to gather additional information before committing to a choice. Working memory mechanisms could serve a role in connecting multiple cues that arrive with staggered timing. This increases the evolutionary pressure for more synergistic strategies.

Moderate synergy ($\lambda = 1.2$ – 1.5) enhances decision accuracy when valid cues co-occur. For instance, this can reflect the combined effect of visual features like colour, size or pattern that are often biologically correlated in valid stimuli such as the presence of the mother hen for chicks [28,46,53]. This model outperforms additive strategies in structured environments, where multiple cues co-occur in target stimuli. When co-occurring cues have low base probability (are rare in the background) and are assigned high bias strength ($\beta > 1$), their combined presence significantly improves performance, even as the number of misleading cues increases (lighter colour lines in the plots). This advantage particularly applies to environments where some misleading cues are present but not overwhelmingly numerous. Such a situation reflects scenarios in which newborns experience a moderately predictable world, as newly hatched fowls in a brood where other chicks and a mother hen are present, and there is a moderate, not overwhelming number of misleading cues.

Strong synergy models have high λ values ($\lambda > 1.5$), reflecting the information advantage carried by cues that are very unlikely to co-occur in the absence of target cues. This depicts the convergence of cues across independent sensory modalities, such as visual and acoustic inputs or tactile and visual signals, that come from independent sources of information. Because such co-occurrences are unlikely to arise by chance, their integration provides a reliable evidence of stimulus validity as target predictor. This model maintains high decision accuracy even in the presence of multiple misleading cues, provided the valid cues span distinct sensory channels. High synergy appears to be present in chicks, where tactile experiences acquired during hatching influence subsequent visual preferences [89], and evidence of visuo-acoustic cross-modal matching [90].

Overall, this model provides a framework for explaining how decision success in naive organisms can be shaped by the number and extrinsic probability of cues, but also by intrinsic biases and synergy. Moreover, the quantitative nature of the model means that it can be used to make specific hypotheses about cue integration for future experiments and for secondary data analysis. While this simple model can explain some decision-making dynamics, we are aware that it does not capture the full complexity behind neonate behavioural responses.

The empirical values of parameters can be determined based on the strength of individual biases tested in isolation, base probabilities of features encountered in specific environments and synergistic effects estimated from the strength of individual

biases tested when multiple cues are present at the same time. Multivariate interaction parameters can be inferred from experimental data by different methods, such as regression or Bayesian inference. We also find value in exploring the general qualitative behaviour of the model with generic parameter values not specifically matched to data.

6. Discussion

We reviewed literature on decision-making in neonate animals and developed a model to understand how inexperienced animals can benefit from innate biases as guides to early-life choices, while reducing the chances of making wrong decisions. Based on the several early predispositions observed in newborn animals [2,18,27], we modelled innate (not learned) preferences for individual cues as soft, probabilistic priors rather than rigid, fixed-action patterns. We put forward a quantitative model of how multiple weak cues could be integrated, accommodating intrinsic bias, cue rarity, misleading cues and cue independence, to reduce the chances that an organism makes a maladaptive response.

We showed that an effective strategy in the evolution of biases/predisposition is focusing on the adaptive rare and co-occurring features present in target objects, since this reduces false positives. Moreover, the presence of multiple weak biases for cues expected in target stimuli can effectively guide behavioural choices as a self-supervised strategy of decision-making that leverages the intrinsic structure of the environment to increase the adaptive value of spontaneous, innate choices that do not depend on learning. While neuroscience and machine learning [91,92] research has put forward self-supervised strategies as a learning strategy that leads experience-dependent development of the central nervous system (reviewed in [93]), our model for the first time applies this concept to decision-making in naive animals, before any learning has taken place. Although the evolutionary processes that shape innate biases have sometimes been framed by analogy as a form of slow learning (see artificial intelligence literature [16,94]), there is a difference between predisposed mechanisms that are shaped by evolutionary mechanisms and do not depend on specific experience, and learning, which depends on experience and changes within the lifetime of an individual. Thus, our model applies to all inexperienced systems, including artificial neural networks, whose architecture and hyperparameters can implement inductive biases akin to what we have modelled.

We focus in particular on soft spontaneous preferences: probabilistic priors rather than the rigid fixed-action patterns emphasized by early ethologists [19,20,95]. This model reflects recent empirical evidence showing that early predispositions [2,18,27,47] are not only widespread across species [1,43,48,59], but also transient [77,78], and variable between individuals [36,55]. For instance, predisposed preferences for colours can be modified by simple exposure to stimuli initially not preferred [36,80], and preferences for female chicks [79] or for some motion dynamics [77,78] spontaneously fade away outside specific time windows, and different individuals might exhibit opposite preferences compared with the overall population average [36,37,50,55,56,61]. The use of multiple weak biases as an adaptive strategy for decision-making in the absence of previous experience can explain why multiple biases are observed and why these biases are widespread across species [1,43,48,96], while at the same time being weak when considered in isolation. Hence our model can explain part of the observed complexity.

Our model accommodates additive and synergistic cue integration strategies (via the continuous parameter λ) and variable ratios of valid and misleading cues. Synergistic integration leverages the co-occurrence of cues to amplify behavioural responses in response of co-occurring evidence. The advantage of using multiple cues and synergistic integration increases in complex environments with more misleading cues. This scenario is consistent with empirical findings: co-occurring cues facilitate behavioural responses in newborn chicks [51,81,82], the co-occurrence of multiple independent cues—such as olfactory and visual, or visual and acoustic cues—strengthens biased choices [4], visual preferences in chicks can be modulated by corresponding features presented in different modalities (e.g. cross-modal tactile to visual recognition at first sight [89]). In general, cross-modal cues are particularly promising for synergistic integration, since they are more likely to be uncorrelated amidst random, misleading cues. By scaling each cue by a measure of its rarity in the environment, the model boosts the default weight of cues that are rare in the environment (e.g. red and blue hue [36,51], or animate movement dynamics [37,55,56]). The very rarity of such cues in background environments increases the likelihood that they represent adaptive stimuli. In combination with the model's bias and synergy terms, these parameters flexibly capture the space of strategies organisms can use to optimize early decisions, even in the absence of prior learning.

Overall, our model offers a conceptual framework to understand the presence of numerous weak biases observed in newborn animals, before learning has occurred. Rather than triggering automatic responses to isolated stimuli, synergistic interactions among multiple biases can serve as an effective self-supervised mechanism for adaptive, rapid decision-making in complex environments. Synergistic integration of weak biases may also improve the development of AI architectures that might require less or no supervision in decision-making. Previous work on precocial animals has shown that spontaneous biases enhance learning [36]. This result suggests that studying how multiple weak biases support adaptive decision-making in different contexts is a promising avenue for understanding how natural and artificial decision-making systems can efficiently make adaptive decisions, with scalable architectures. Our model provides a suite of testable, quantitative predictions about cue integration that can be used to test these ideas.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

Data accessibility. Supplementary material is available online [97].

Declaration of AI use. We used AI-assisted technologies as a search engine and to improve the readability of the work.

Authors' contributions. E.V.: conceptualization, formal analysis, funding acquisition, methodology, project administration, software, validation, visualization, writing—original draft, writing—review and editing; B.d.B.: conceptualization, funding acquisition, methodology, validation, visualization, writing—review and editing.

Both authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. E.V. has been supported by the Leverhulme Trust research grant RPG-2020-287, the Royal Society Leverhulme Trust fellowship SRF\R1\21000155, the BBSRC grant BB/S003223/1. B.d.B. was funded by the prematurely terminated and then reinstated grant NIH/NINDS 1R01NS121874-0.

Acknowledgements. We thank Gabriel Kreiman (Harvard University), Davide Zoccolan (SISSA) and Richard Nichols (Queen Mary University of London) for fruitful discussions on this project.

References

- Versace E, Damini S, Stancher G. 2020 Early preference for face-like stimuli in solitary species as revealed by tortoise hatchlings. *Proc. Natl Acad. Sci. USA* **117**, 24047–24049. (doi:10.1073/pnas.2011453117)
- Rosa-Salva O, Mayer U, Versace E, Hébert M, Lemaire BS, Vallortigara G. 2021 Sensitive periods for social development: interactions between predisposed and learned mechanisms. *Cognition* **213**, 104552. (doi:10.1016/j.cognition.2020.104552)
- Hébert M, Versace E, Vallortigara G. 2019 Inexperienced preys know when to flee or to freeze in front of a threat. *Proc. Natl Acad. Sci. USA* **116**, 22918–22920. (doi:10.1073/pnas.1915504116)
- Mishra A, Jain A, Iyer PS, Suryanarayanan A, Nordström K, Olsson SB. 2025 Innate floral object identification in a solitary pollinator employs a combination of both visual and olfactory cues. *Sci. Nat.* **112**, 18. (doi:10.1007/s00114-025-01965-5)
- Protti-Sánchez F, Mayer U, Rowland HM. 2023 In paired preference tests, domestic chicks innately choose the colour green over red, and the shape of a frog over a sphere when both stimuli are green. *Anim. Cogn.* **26**, 1973–1983. (doi:10.1007/s10071-023-01821-x)
- Lieder F, Griffiths TL. 2015 When to use which heuristic: a rational solution to the strategy selection problem. In *Proc. of the 37th Annual Conf. of the Cognitive Science Society*, p. 1367. Cognitive Science Society. <https://cogsci.mindmodeling.org/2015/papers/0239/index.html>.
- Greggor AL, Clayton NS, Phalan B, Thornton A. 2014 Comparative cognition for conservationists. *Trends Ecol. Evol.* **29**, 489–495. (doi:10.1016/j.tree.2014.06.004)
- Koehler DJ, Harvey N (eds). 2009 *Blackwell handbook of judgment and decision making*. Malden, MA, USA: Blackwell. (doi:10.1002/9780470752937)
- Gwynne DT, Rentz DCF. 1983 Beetles on the bottle: male buprestids mistake stubbies for females (Coleoptera). *Aust. J. Entomol.* **22**, 79–80. (doi:10.1111/j.1440-6055.1983.tb01846.x)
- Greggor AL, Berger-Tal O, Blumstein DT. 2020 The rules of attraction: the necessary role of animal cognition in explaining conservation failures and successes. *Annu. Rev. Ecol. Evol. Syst.* **51**, 483–503. (doi:10.1146/annurev-ecolsys-011720-103212)
- Truscott Z, Booth DT, Limpus CJ. 2017 The effect of on-shore light pollution on sea-turtle hatchlings commencing their off-shore swim. *Wildl. Res.* **44**, 127. (doi:10.1071/wr16143)
- Versace E. 2022 Precocial. In *Encyclopedia of animal cognition and behavior* (eds J Vonk, TK Shackelford), pp. 5488–5490. Cham, Switzerland: Springer International Publishing. (doi:10.1007/978-3-319-55065-7_459)
- Ralls K, Lundrigan B, Kranz K. 1987 Mother-young relationships in captive ungulates: behavioral changes over time. *Ethology* **75**, 1–14. (doi:10.1111/j.1439-0310.1987.tb00637.x)
- Goyal A, Bengio Y. 2022 Inductive biases for deep learning of higher-level cognition. *Proc. R. Soc. A* **478**, 20210068. (doi:10.1098/rspa.2021.0068)
- Zhang R, Pitkow X, Angelaki DE. 2024 Inductive biases of neural network modularity in spatial navigation. *Sci. Adv.* **10**, k1256. (doi:10.1126/sciadv.adk1256)
- Botvinick M, Ritter S, Wang JX, Kurth-Nelson Z, Blundell C, Hassabis D. 2019 Reinforcement learning, fast and slow. *Trends Cogn. Sci.* **23**, 408–422. (doi:10.1016/j.tics.2019.02.006)
- Versace E, Vallortigara G. 2015 Origins of knowledge: insights from precocial species. *Front. Behav. Neurosci.* **9**, 338. (doi:10.3389/fnbeh.2015.00338)
- Versace E, Martinho-Truswell A, Kacelnik A, Vallortigara G. 2018 Priors in animal and artificial intelligence: where does learning begin? *Trends Cogn. Sci.* **22**, 963–965. (doi:10.1016/j.tics.2018.07.005)
- Tinbergen N. 1952 The curious behavior of the stickleback. *Sci. Am.* **187**, 22–26. (doi:10.1038/scientificamerican1252-22)
- Tinbergen N. 2018 *The study of instinct*, 1st ed. reprinted. New York, NY, USA: Oxford University Press.
- Mederos S, Blakely P, Vissers N, Clopath C, Hofer SB. 2025 Overwriting an instinct: visual cortex instructs learning to suppress fear responses. *Science* **387**, 682–688. (doi:10.1126/science.adr2247)
- De Franceschi G, Vivattanasarn T, Saleem AB, Solomon SG. 2016 Vision guides selection of freeze or flight defense strategies in mice. *Curr. Biol.* **26**, 2150–2154. (doi:10.1016/j.cub.2016.06.006)
- Buckner C. 2018 Empiricism without magic: transformational abstraction in deep convolutional neural networks. *Synthese* **195**, 5339–5372. (doi:10.1007/s11229-018-01949-1)
- Wood JN. 2024 Artificial intelligence tackles the nature–nurture debate. *Nat. Mach. Intell.* **6**, 381–382. (doi:10.1038/s42256-024-00828-4)
- Orhan AE, Lake BM. 2024 Learning high-level visual representations from a child's perspective without strong inductive biases. *Nat. Mach. Intell.* **6**, 271–283. (doi:10.1038/s42256-024-00802-0)
- Pandey L, Lee D, Wood SMW, Wood JN. 2024 Parallel development of object recognition in newborn chicks and deep neural networks. *PLoS Comput. Biol.* **20**, e1012600. (doi:10.1371/journal.pcbi.1012600)
- Lemaire BS, Vallortigara G. 2022 Life is in motion (through a chick's eye). *Anim. Cogn.* **26**, 129–140. (doi:10.1007/s10071-022-01703-8)
- Johnson MH, Horn G. 1988 Development of filial preferences in dark-reared chicks. *Anim. Behav.* **36**, 675–683. (doi:10.1016/s0003-3472(88)80150-7)
- Evans DA, Stempel AV, Vale R, Branco T. 2019 Cognitive control of escape behaviour. *Trends Cogn. Sci.* **23**, 334–348. (doi:10.1016/j.tics.2019.01.012)
- Fotowat H, Gabbiani F. 2011 Collision detection as a model for sensory-motor integration. *Annu. Rev. Neurosci.* **34**, 1–19. (doi:10.1146/annurev-neuro-061010-113632)
- Krizhevsky A, Sutskever I, Hinton G. 2012 ImageNet classification with deep convolutional neural networks. *Commun. ACM* **60**, 84–90. (doi:10.1145/3065386)
- Hibi SE, Bengio Y. 1995 Hierarchical recurrent neural networks for long-term dependencies. In *NIPS'95: Proc. of the 9th Int. Conf. on Neural Information Processing Systems*, pp. 493–494. ACM. (doi:10.5555/2998828.2998898)
- Wheatcroft D, Qvarnström A. 2015 A blueprint for vocal learning: auditory predispositions from brains to genomes. *Biol. Lett.* **11**, 20150155. (doi:10.1098/rsbl.2015.0155)
- Zatorre RJ. 2013 Predispositions and plasticity in music and speech learning: neural correlates and implications. *Science* **342**, 585–589. (doi:10.1126/science.1238414)
- Olszewska AM, Gaca M, Herman AM, Jednoróg K, Marchewka A. 2021 How musical training shapes the adult brain: predispositions and neuroplasticity. *Front. Neurosci.* **15**, 630829. (doi:10.3389/fnins.2021.630829)
- Wang S, Vasas V, Freeland L, Osorio D, Versace E. 2024 Spontaneous biases enhance generalization in the neonate brain. *iScience* **27**, 110195. (doi:10.1016/j.isci.2024.110195)

37. Bliss L, Vasas V, Freeland L, Roach R, Ferrè ER, Versace E. 2023 A spontaneous gravity prior: newborn chicks prefer stimuli that move against gravity. *Biol. Lett.* **19**, 20220502. (doi:10.1098/rsbl.2022.0502)
38. Vallortigara G, Versace E. 2022 Filial Imprinting. In *Encyclopedia of animal cognition and behavior*, pp. 1943–1948. Cham, Switzerland: Springer. (doi:10.1007/978-3-319-55065-7_1989)
39. Bateson P. 1966 The characteristics and context of imprinting. *Biol. Rev. Camb. Philos. Soc.* **41**, 177–211. (doi:10.1111/j.1469-185x.1966.tb01489.x)
40. McCabe BJ. 2013 Imprinting. *Wiley Interdiscip. Rev.* **4**, 375–390. (doi:10.1002/wcs.1231)
41. Bolhuis JJ. 1991 Mechanisms of avian imprinting: a review. *Biol. Rev.* **66**, 303–345. (doi:10.1111/j.1469-185x.1991.tb01145.x)
42. Nicol CJ. 2015 In *The behavioural biology of chickens*. Oxfordshire, UK: CABI. (doi:10.1079/9781780642499.0000)
43. Kobylkov D, Rosa-Salva O, Zanon M, Vallortigara G. 2024 Innate face-selectivity in the brain of young domestic chicks. *Proc. Natl Acad. Sci. USA* **121**, e2410404121. (doi:10.1073/pnas.2410404121)
44. Horn G, McCabe BJ. 1984 Predispositions and preferences: effects on imprinting of lesions to the chick brain. *Anim. Behav.* **32**, 288–292. (doi:10.1016/s0003-3472(84)80349-8)
45. Mayer U, Rosa-Salva O, Lorenzi E, Vallortigara G. 2016 Social predisposition dependent neuronal activity in the intermediate medial mesopallium of domestic chicks (*Gallus gallus domesticus*). *Behav. Brain Res.* **310**, 93–102. (doi:10.1016/j.bbr.2016.05.019)
46. Mayer U, Rosa-Salva O, Loveland JL, Vallortigara G. 2019 Selective response of the nucleus taeniae of the amygdala to a naturalistic social stimulus in visually naive domestic chicks. *Sci. Rep.* **9**, 9849. (doi:10.1038/s41598-019-46322-5)
47. Kobylkov D, Vallortigara G. 2024 Face detection mechanisms: nature vs. nurture. *Front. Neurosci.* **18**, 1404174. (doi:10.3389/fnins.2024.1404174)
48. Rosa-Salva O, Regolin L, Vallortigara G. 2010 Faces are special for newly hatched chicks: evidence for inborn domain-specific mechanisms underlying spontaneous preferences for face-like stimuli. *Dev. Sci.* **13**, 565–577. (doi:10.1111/j.1467-7687.2009.00914.x)
49. Rosa-Salva O, Mayer U, Vallortigara G. 2019 Unlearned visual preferences for the head region in domestic chicks. *PLoS One* **14**, e0222079. (doi:10.1371/journal.pone.0222079)
50. Versace E, Schill J, Nencini AM, Vallortigara G. 2016 Naïve chicks prefer hollow objects. *PLoS One* **11**, e0166425. (doi:10.1371/journal.pone.0166425)
51. Freeland LV, Emmerson MG, Vasas V, Gomes J, Versace E. 2025 Assessing preferences for adult versus juvenile features in young animals: Newly hatched chicks spontaneously approach red and large stimuli. *Learn. Behav.* **53**, 145–156. (doi:10.3758/s13420-024-00638-z)
52. Kovach JK. 1971 Effectiveness of different colors in the elicitation and development of approach behavior in chicks. *Behaviour* **38**, 154–168. (doi:10.1163/156853971x00069)
53. Versace E, Fracasso I, Baldan G, Dalle Zotte A, Vallortigara G. 2017 Newborn chicks show inherited variability in early social predispositions for hen-like stimuli. *Sci. Rep.* **7**, 40296. (doi:10.1038/srep40296)
54. Schulman AH, Hale EE, Graves HB. 1970 Visual stimulus characteristics for initial approach response in chicks (*Gallus domesticus*). *Anim. Behav.* **18**, 461–466. (doi:10.1016/0003-3472(70)90040-0)
55. Vallortigara G, Regolin L, Marconato F. 2005 Visually inexperienced chicks exhibit spontaneous preference for biological motion patterns. *PLoS Biol.* **3**, e208. (doi:10.1371/journal.pbio.0030208)
56. Rosa-Salva O, Hernik M, Broseghini A, Vallortigara G. 2018 Visually-naïve chicks prefer agents that move as if constrained by a bilateral body-plan. *Cognition* **173**, 106–114. (doi:10.1016/j.cognition.2018.01.004)
57. Mascalcioni E, Regolin L, Vallortigara G. 2010 Innate sensitivity for self-propelled causal agency in newly hatched chicks. *Proc. Natl Acad. Sci. USA* **107**, 4483–4485. (doi:10.1073/pnas.0908792107)
58. Di Giorgio E, Lunghi M, Simion F, Vallortigara G. 2016 Visual cues of motion that trigger animacy perception at birth: the case of self-propulsion. *Dev. Sci.* **20**, 1–12. (doi:10.1111/desc.12394)
59. Viola Macchi C, Turati C, Simion F. 2004 Can a nonspecific bias toward top-heavy patterns explain newborns' face preference? *Psychol. Sci.* **15**, 379–383. (doi:10.1111/j.0956-7976.2004.00688.x)
60. Buiatti M, Di Giorgio E, Piazza M, Polloni C, Menna G, Taddei F, Baldo E, Vallortigara G. 2019 Cortical route for facelike pattern processing in human newborns. *Proc. Natl Acad. Sci. USA* **116**, 4625–4630. (doi:10.1073/pnas.1812419116)
61. Bardi L, Regolin L, Simion F. 2011 Biological motion preference in humans at birth: role of dynamic and configural properties. *Dev. Sci.* **14**, 353–359. (doi:10.1111/j.1467-7687.2010.00985.x)
62. Hernik M, Fearon P, Csibra G. 2014 Action anticipation in human infants reveals assumptions about anteroposterior body-structure and action. *Proc. R. Soc. B* **281**, 20133205. (doi:10.1098/rspb.2013.3205)
63. Matsushima T, Izumi T, Vallortigara G. 2024 The domestic chick as an animal model of autism spectrum disorder: building adaptive social perceptions through prenatally formed predispositions. *Front. Neurosci.* **18**, 1279947. (doi:10.3389/fnins.2024.1279947)
64. Krebs JR, Davies NB (eds). 1984 *Behavioural ecology: an evolutionary approach*, 2nd edn. Sunderland, MA, USA: Sinauer Associates.
65. Herczeg G, Ghani NIA, Merilä J. 2016 On plasticity of aggression: influence of past and present predation risk, social environment and sex. *Behav. Ecol. Sociobiol.* **70**, 179–187. (doi:10.1007/s00265-015-2037-3)
66. Lenzi SC, Cossell L, Grainger B, Olesen SF, Branco T, Margrie TW. 2022 Threat history controls flexible escape behavior in mice. *Curr. Biol.* **32**, 2972–2979. (doi:10.1016/j.cub.2022.05.022)
67. Ache JM, Namiki S, Lee A, Branson K, Card GM. 2019 State-dependent decoupling of sensory and motor circuits underlies behavioral flexibility in *Drosophila*. *Nat. Neurosci.* **22**, 1132–1139. (doi:10.1038/s41593-019-0413-4)
68. Bekoff A. 1999 Development of motor patterns in avian embryos: control of hatching behaviour. In *Proc. of the 22nd Int. Ornithological Congress*, Durban, South Africa, pp. 2776–2783. Johannesburg, South Africa: BirdLife South Africa.
69. Kain JS, Stokes C, de Bivort BL. 2012 Phototactic personality in fruit flies and its suppression by serotonin and white. *Proc. Natl Acad. Sci. USA* **109**, 19834–19839. (doi:10.1073/pnas.1211988109)
70. Hawkins, Magurran, Armstrong. 2004 Innate predator recognition in newly-hatched Atlantic salmon. *Behaviour* **141**, 1249–1262. (doi:10.1163/1568539042729694)
71. Ewert JP. 1987 Neuroethology of releasing mechanisms: prey-catching in toads. *Behav. Brain Sci.* **10**, 337–368. (doi:10.1017/s0140525x00023128)
72. Ewert JP. 2004 Motion perception shapes the visual world of amphibians. In *Complex worlds from simpler nervous systems* (ed. FR Prete), pp. 117–160. Cambridge, MA, USA: The MIT Press. (doi:10.7551/mitpress/1994.003.0011)
73. Dunlap AS, Stephens DW. 2016 Reliability, uncertainty, and costs in the evolution of animal learning. *Curr. Opin. Behav. Sci.* **12**, 73–79. (doi:10.1016/j.cobeha.2016.09.010)
74. Mery F, Kawecki TJ. 2002 Experimental evolution of learning ability in fruit flies. *Proc. Natl Acad. Sci. USA* **99**, 14274–14279. (doi:10.1073/pnas.222371199)

75. Rosa Salva O, Mayer U, Vallortigara G. 2015 Roots of a social brain: developmental models of emerging animacy-detection mechanisms. *Neurosci. Biobehav. Rev.* **50**, 150–168. (doi:10.1016/j.neubiorev.2014.12.015)
76. Egorova OV, Anokhin KV. 2003 Experimental analysis of the processes of systems genesis: expression of the c-fos gene in the chick brain during treatments inducing the development of the species-specific results-of-action acceptor. *Neurosci. Behav. Physiol.* **33**, 209–216. (doi:10.1023/a:1022186911789)
77. Versace E, Ragusa M, Vallortigara G. 2019 A transient time window for early predispositions in newborn chicks. *Sci. Rep.* **9**, 18767. (doi:10.1038/s41598-019-55255-y)
78. Lorenzi E, Lemaire BS, Versace E, Matsushima T, Vallortigara G. 2021 Resurgence of an inborn attraction for animate objects via thyroid hormone T3. *Front. Behav. Neurosci.* **15**, 1–8. (doi:10.3389/fnbeh.2021.675994)
79. Pallante V, Rucco D, Versace E. 2021 Young chicks quickly lose their spontaneous preference to aggregate with females. *Behav. Ecol. Sociobiol.* **75**, 78. (doi:10.1007/s00265-021-03012-5)
80. Lemaire BS, Rucco D, Josserand M, Vallortigara G, Versace E. 2021 Stability and individual variability of social attachment in imprinting. *Sci. Rep.* **11**, 7914. (doi:10.1038/s41598-021-86989-3)
81. Miura M, Matsushima T. 2016 Biological motion facilitates filial imprinting. *Anim. Behav.* **116**, 171–180. (doi:10.1016/j.anbehav.2016.03.025)
82. Miura M, Nishi D, Matsushima T. 2019 Combined predisposed preferences for colour and biological motion make robust development of social attachment through imprinting. *Anim. Cogn.* **23**, 169–188. (doi:10.1007/s10071-019-01327-5)
83. Hess EH. 1956 Natural preferences of chicks and ducklings for objects of different colors. *Psychol. Rep.* **2**, 477–483. (doi:10.2466/pr0.1956.2.3.477)
84. Versace E, Ragusa M, Pallante V, Wang S. 2021 Attraction for familiar conspecifics in young chicks (*Gallus gallus*): an interbreed study. *Behav. Process.* **193**, 104498. (doi:10.1016/j.beproc.2021.104498)
85. Renoult JP, Schaefer HM, Sallé B, Charpentier MJE. 2011 The evolution of the multicoloured face of mandrills: insights from the perceptual space of colour vision. *PLoS One* **6**, e29117. (doi:10.1371/journal.pone.0029117)
86. Emberts Z, Wiens JJ. 2022 Why are animals conspicuously colored? Evolution of sexual versus warning signals in land vertebrates. *Evolution* 2879–2891. (doi:10.1111/evo.14636)
87. Mayer U, Rosa-Salva O, Morbioli F, Vallortigara G. 2017 The motion of a living conspecific activates septal and preoptic areas in naive domestic chicks (*Gallus gallus*). *Eur. J. Neurosci.* **45**, 423–432. (doi:10.1111/ejn.13484)
88. Shannon CE. 1948 A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423. (doi:10.1002/j.1538-7305.1948.tb01338.x)
89. Versace E, Freeland L, Emmerson MG. 2024 First-sight recognition of touched objects shows that chicks can solve Molyneux's problem. *Biol. Lett.* **20** (4), 20240025. (doi:10.1098/rsbl.2024.0025)
90. Loconsole M, Benavides-Varela S, Regolin L. 2024 Matching sounds to shapes: evidence of the Bouba-Kiki effect in naïve baby chicks. *bioRxiv* 2024.05.17.594640. (doi:10.1101/2024.05.17.594640)
91. Gui J, Chen T, Zhang J, Cao Q, Sun Z, Luo H, Tao D. 2024 A survey on self-supervised learning: algorithms, applications, and future trends. *arXiv* 2301.05712. (doi:10.48550/arXiv.2301.05712)
92. Ororbias AG, Friston K, Rao RPN. 2025 Meta-representational predictive coding: biomimetic self-supervised learning. *arXiv* 2503.21796. (doi:10.48550/arXiv.2503.21796)
93. Matteucci G, Piasini E, Zoccolan D. 2024 Unsupervised learning of mid-level visual representations. *Curr. Opin. Neurobiol.* **84**, 102834. (doi:10.1016/j.conb.2023.102834)
94. Hasson U, Nastase SA, Goldstein A. 2020 Direct fit to nature: an evolutionary perspective on biological and artificial neural networks. *Neuron* **105**, 416–434. (doi:10.1016/j.neuron.2019.12.002)
95. Lorenz K, Tinbergen N. 1939 Taxis und Instinkthandlung in der Eirollbewegung der Graugans. I. *Z. Für Tierpsychol.* **2**, 1–29. (doi:10.1111/j.1439-0310.1939.tb01558.x)
96. Simion F, Di Giorgio E. 2015 Face perception and processing in early infancy: inborn predispositions and developmental changes. *Front. Psychol.* **6**, 969. (doi:10.3389/fpsyg.2015.00969)
97. Versace E, de Bivort B. 2026 Supplementary material from: Multiple weak biases support adaptive choices without prior experience: a self-supervised strategy. Figshare. (doi:10.6084/m9.figshare.c.8230302)