PROCEEDINGS B

rspb.royalsocietypublishing.org

Research

Cite this article: Panchanathan K, Frankenhuis WE. 2016 The evolution of sensitive periods in a model of incremental development. *Proc. R. Soc. B* **283**: 20152439. http://dx.doi.org/10.1098/rspb.2015.2439

Received: 13 October 2015
Accepted: 21 December 2015

Subject Areas:
evolution, theoretical biology, developmental biology

Keywords:
sensitive periods, evolution, development, adaptation, cues, Bayesian updating

Author for correspondence:
Willem E. Frankenhuis
e-mail: wfrankenhuis@gmail.com

†These authors contributed equally to this study.

Electronic supplementary material is available at http://dx.doi.org/10.1098/rspb.2015.2439 or via http://rspb.royalsocietypublishing.org.

THE ROYAL SOCIETY PUBLISHING

The evolution of sensitive periods in a model of incremental development

Karthik Panchanathan[1,†] and Willem E. Frankenhuis[2,†]

[1]Department of Anthropology, University of Missouri, 107 Swallow Hall, Columbia, MO 65211-1440, USA
[2]Behavioural Science Institute, Radboud University Nijmegen, Montessorilaan 3, PO Box 9104, Nijmegen, 6500 HE, The Netherlands

Sensitive periods, in which experience shapes phenotypic development to a larger extent than other periods, are widespread in nature. Despite a recent focus on neural–physiological explanation, few formal models have examined the evolutionary selection pressures that result in developmental mechanisms that produce sensitive periods. Here, we present such a model. We model development as a specialization process during which individuals incrementally adapt to local environmental conditions, while receiving a constant stream of cost-free, imperfect cues to the environmental state. We compute optimal developmental programmes across a range of ecological conditions and use these programmes to simulate developmental trajectories and obtain distributions of mature phenotypes. We highlight four main results. First, matching the empirical record, sensitive periods often result from experience or from a combination of age and experience, but rarely from age alone. Second, individual differences in sensitive periods emerge as a result of stochasticity in cues: individuals who obtain more consistent cue sets lose their plasticity at faster rates. Third, in some cases, experience shapes phenotypes only at a later life stage (lagged effects). Fourth, individuals might perseverate along developmental trajectories despite accumulating evidence suggesting the alternate trajectory is more likely to match the ecology.

> . . .we all begin with the natural equipment to live a thousand kinds of life but end in the end having lived only one.
>
> —Clifford Geertz, 1973

1. Introduction

Phenotypic plasticity, the ability to tailor development to local conditions, is widespread in nature [1] and evolves in environments that vary spatially or temporally if cues provide information about the state of the environment [2,3]. The degree of plasticity, however, is not fixed: it varies between species, between individuals, between traits within an individual, and within traits across development. Recent neural–physiological research offers mechanistic explanations of such variation [4–12]. Despite theoretical progress explaining the evolution of sensitive periods (e.g. [13–19]), there are few formal models (for exceptions, see [20–26]). Here, we present such a model.

We conceptualize development as a specialization process during which individuals simultaneously learn about and incrementally adapt to local environmental conditions [21,27]. This approach differs from existing models, which typically assume a two-stage life history, in which organisms first obtain environmental cues and later develop phenotypes, either immediately (e.g. [28,29]) or after a time lag (e.g. [26,30]). Trait development poses special challenges when earlier onset of specialization has the potential for higher fitness but also increases the likelihood of phenotype–environment mismatch. For instance, if trait construction takes time, or if earlier integration of phenotypic components increases the efficiency of traits ('the epiphenotype problem'; [31]), organisms benefit from early onset of trait development (e.g. anti-predator defences; [32]). But, an early commitment increases the risk of phenotype–environment mismatch [33], due to fewer opportunities to collect cues, which improve estimates of the environmental state [34–36].

In a previous model [21], we showed that sensitive periods are likely to evolve if organisms face a trade-off between sampling cues and phenotypic specialization. In that scenario, organisms often sample cues early in ontogeny. Once phenotypic construction begins, however, plasticity terminates as a consequence of the assumed trade-off. Here, we present a novel model in which learning and specialization do not trade-off. Instead, individuals receive cues throughout ontogeny while building their phenotypes. This allows organisms to revise their estimates about the environmental state and switch developmental trajectories.

This model makes several other assumptions. First, fitness increases monotonically with the time invested in correctly tailoring the phenotype to local conditions. Accordingly, organisms benefit from an earlier onset of specialization, allowing more time for refining the phenotype–environment fit [21]. For example, water fleas that begin tailoring their phenotype prenatally (inside the mother) towards a predator-rich environment develop larger protective helmets than conspecifics who start doing so only after hatching [37].

Second, developing a phenotype that does not match the environmental state is costly, and these costs depend on the degree of mismatch. In our model, organisms can specialize towards one of two phenotypes. With no mismatch costs, the two phenotype extremes should be conceptualized as two orthogonal and non-antagonistic dimensions. With mismatch costs, the two phenotype extremes are orthogonal and antagonistic. By antagonistic we mean a phenotypic increment increases an organism's fitness if it is matched to the environment and reduces an organism's fitness if it is mismatched. The two phenotype extremes should not be interpreted as ends of a single dimension (e.g. pace of life-history development).

Third, organisms are uncertain about the environmental state but receive cost-free, imperfect cues that improve their estimates. Our previous model [21] assumed that learning about the environment trades off with specializing towards a phenotypic target. This represents 'costly' information, which includes active information search (e.g. barnacles inspecting different rocks before settling). The current model assumes no trade-off between learning and specialization. This represents 'cost-free' information, which includes passive information reception (e.g. water fleas detecting predators while foraging).

We use state-dependent life-history theory [38–40], implemented by stochastic dynamic programming, to model how natural selection shapes developmental mechanisms that learn about local conditions by receiving a constant stream of cost-free but imperfect cues, and incrementally adapt to these conditions, with correct increments increasing fitness and incorrect increments decreasing fitness.

2. Model

(a) The organism and the environment

The environment is structured into an infinite number of discrete patches. Each patch is in one of two states: E_1 or E_0 (e.g. dangerous or safe). The state of a patch is stable within generations, and the distribution of patches remains stable across generations (i.e. the environment is characterized by spatial variation, not temporal variation [41]). There is one optimal phenotype for each environmental state: P_1 or P_0 (e.g. having spent all of ontogeny specializing towards an armoured or a sleek phenotype). After birth, organisms disperse to randomly selected patches, receive cues and develop phenotypes, reproduce at a rate proportional to their phenotype–environment fit, die, and the cycle repeats. Because the distribution of patches remains fixed across evolutionary time, we assume that developmental programmes have adapted to this distribution [42].

Ontogeny (i.e. the developmentally relevant stage for the trait of interest; not necessarily early life) consists of 20 discrete and non-overlapping time periods. In each period, organisms receive a cue (C_1 or C_0) and then choose to either increment (1/20th of the way) towards P_1 or P_0, or to wait and forgo phenotypic specialization. Once developed, phenotypic increments cannot be reversed (e.g. discarded or resorbed).

Cues indicate the environmental state (e.g. dangerous or safe) with a fixed probability. The cue validity is the probability of receiving a given cue (e.g. detecting the chemical signature of a predator) conditioned on being in the corresponding environmental state (e.g. dangerous). We assume that the cue validities of each environmental state are equal: $P(C_1|E_1) = P(C_0|E_0)$. The probability of receiving an incorrect cue, $P(C_1|E_0)$ or $P(C_0|E_1)$, is the complement of receiving the correct cue, $1 - P(C_0|E_0)$ or $1 - P(C_1|E_1)$. Organisms learn about the environmental state using Bayes' theorem, the optimal way of information updating [42–44] if updating is cost-free. Organisms use the fixed distribution of patches as the Bayesian prior and the fixed cue validities to update their estimates.

(b) Penalty and reward functions

Organisms comprise myriad traits (e.g. eye colour, wing shape). Different variants of any trait may result in different fitness consequences, deviations from some baseline level. The fitness of an organism increases relative to baseline with each correct phenotypic increment (i.e. the phenotype corresponds to the ecology) and decreases with each incorrect increment (see electronic supplementary material, appendix S6, for the dynamic programming equations, including fitness functions, used to simulate optimal developmental programmes). We consider three functional mappings between correct phenotypic development and fitness: linear rewards (the marginal increase in fitness is constant with each correct increment), diminishing rewards (the marginal increase decreases) and increasing rewards (the marginal increase increases). We consider three analogous mappings between incorrect phenotypic development and fitness decrements (i.e. linear penalties, diminishing penalties, increasing penalties).

We assume that the fitness penalty for a completely mismatched organism is equal to the fitness reward for a completely matched organism (we discuss the consequence of varying the penalty magnitude in the Discussion section). The pay-offs for correct and incorrect specialization are the same in both environmental states. Further, phenotypic specialization does not affect fitness during the life stage in which it is constructed, but at some later life stage. Models of predictive adaptive response (e.g. [32,33]) and time lag (e.g. [26,30]) share this assumption. In our model, however, the adaptive fit later in life depends on the time invested in progressively constructing a phenotype.

(c) Quantifying plasticity

We quantify plasticity using an experimental twin study. Using the optimal developmental policy, we simulate twins who are identical in every way up to time period t, and then keep one twin (the *original*) in its natal patch and send

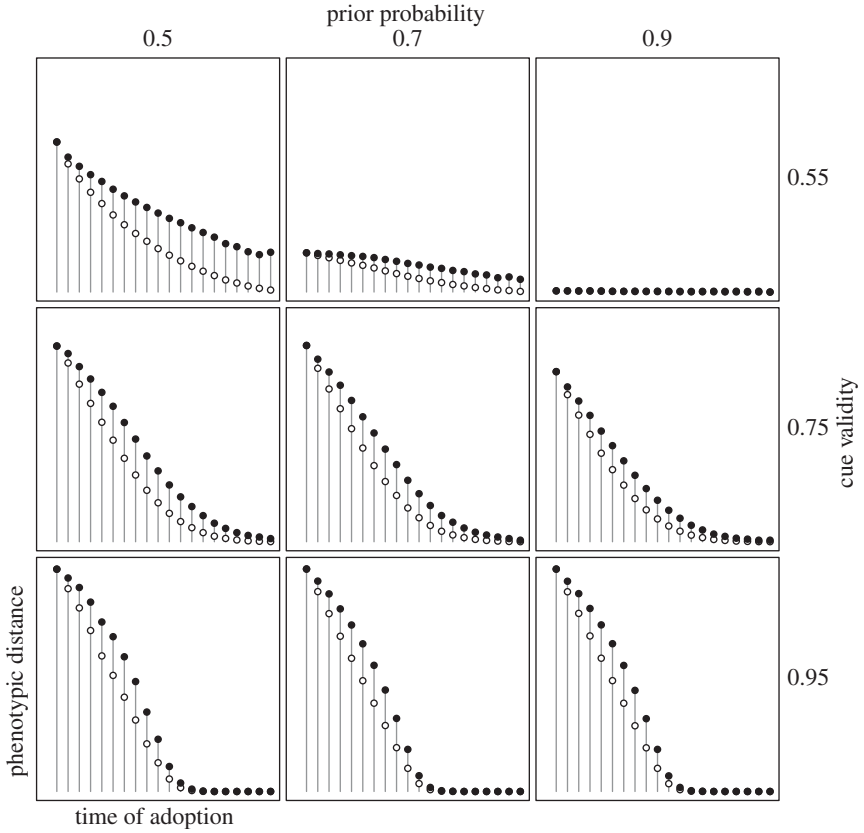

Figure 1. Developmental plasticity. This plot depicts the phenotypic distance between organisms and their doppelgängers with linear rewards for correct phenotypic development and linear penalties for incorrect development. Each row of three panels depicts a cue validity (0.55, 0.75, 0.95) and each column of three panels a prior probability of E_1 (0.5, 0.7, 0.9). Within each panel, the horizontal axis depicts the time period in which the adoption occurred and the vertical axis phenotypic distance. Phenotypes are defined by two numbers: time periods specializing towards P_1 and towards P_0. Phenotypic distance is the Euclidean distance between two organisms. Open circles depict 'absolute' phenotypic distance: the average distance between 10 000 organisms and their doppelgängers at the end of development (ranging from 0 to $20\sqrt{2}$, scaled to a 0 to 1 range). Filled circles depict 'proportional' distance: absolute distance divided by maximum possible distance (ranging from 0 to 1).

the other twin (the *doppelgänger*) to a mirror patch, after which the doppelgänger receives the opposite cues from the original twin (yoked, opposite cues; not cues from an opposite patch) until the end of ontogeny. This approach resembles methods that are commonly used in empirical studies of sensitive periods (e.g. [45,46]).

We replicate this simulation for 10 000 sets of twins. We quantify plasticity by comparing the phenotypic distance between twins at the time of the adoption to the phenotypic distance at the end of ontogeny. The phenotype is characterized by two numbers: the number of time periods an organism specializes towards each of the two optimal phenotypes (these numbers summed plus the number of forgone specializations equals 20, the total number of time periods). We define the phenotypic distance between two individuals as the Euclidean distance along these two dimensions (i.e. the square root of the sum of the squared differences). We compute the 'absolute' phenotypic distance by taking the average phenotypic distance across the 10 000 twin pairs. We then divide the absolute distance by the maximum possible distance to obtain the 'proportional' phenotypic distance. This measure controls for potential distance, which necessarily declines as ontogeny proceeds.

3. Results

We present four kinds of results. First, we show plasticity by comparing originals with their doppelgängers (figure 1 and electronic supplementary material, figures A1.1–A1.8). Second, we compare the optimal developmental policies (i.e. those favoured by natural selection) across a range of ecological conditions: fitness function, penalty function, cue validity and prior probability (figure 2 and electronic supplementary material, figures A2.1–A2.8). Third, we compare the optimal policies, which may include plasticity, with two non-plastic types that ignore cues (figure 3 and electronic supplementary material, figures A3.1–A3.8). Fourth, we simulate developmental trajectories using these optimal policies to generate distributions of mature phenotypes (figure 4 and electronic supplementary material, figures A4.1–A4.8). (In the electronic supplementary material, figures A5.1–A5.9, we also present the expected degree of phenotypic mismatch.) We focus our analysis on linear reward and linear penalty functions in the main text. Results from decreasing and increasing marginal rewards and penalties appear in the electronic supplementary material.

(a) Plasticity across ontogeny

We first consider the extent to which doppelgängers phenotypically diverge from their original counterparts due to receiving opposite cues after adoption (figure 1 and electronic supplementary material, figures A1.1–A1.8). Results show that the later adoptions occur, the less doppelgängers and their original counterparts diverge (i.e. plasticity declines across ontogeny). The more informative the cues are, the faster plasticity declines, regardless of the prior probability.

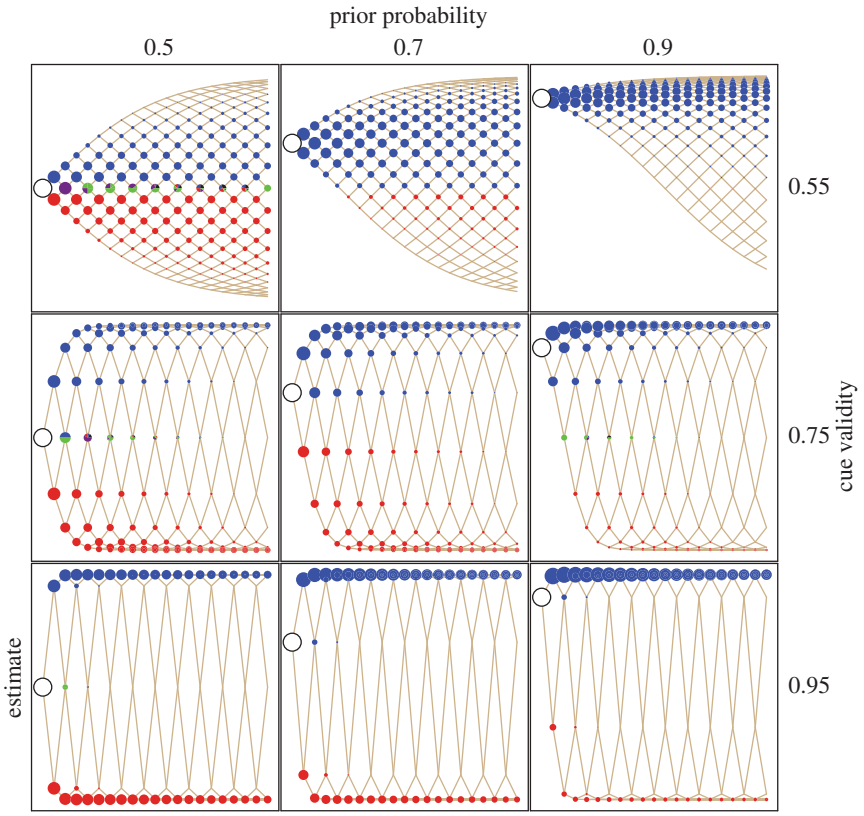

Figure 2. Optimal developmental programmes with linear rewards for correct phenotypic development and linear penalties for incorrect development. Each row of three panels depicts a cue validity (0.55, 0.75, 0.95) and each column of three panels a prior probability of E_1 (0.5, 0.7, 0.9). Within each panel, the horizontal axis depicts developmental time and the vertical axis the organism's estimate of being in E_1. All organisms start with the same estimate (the large white circle). In each time period, organisms sample cues, update estimates and make developmental decisions. Beige lines represent possible estimates across development. Decisions are depicted by coloured circles: black represents waiting, blue specializing towards P_1, red specializing towards P_0, purple a random choice between specializing towards P_1 and P_0, and green a random choice between waiting and specializing towards P_1 and/or P_0. Pies depict situations in which organisms with the same estimate make different decisions. The area of a circle is proportional to the probability of reaching a particular state. Within a time period, these probabilities sum to one.

When cues are informative (e.g. 0.7 or 0.9), plasticity often drops to zero—in some cases well before the end of the ontogeny. More accurate cues imply more consistent cue sets, resulting in more certain posteriors (at the time of adoption). The more certain an organism is, the more counter-evidence it will need to revise its estimate (to the point where switching phenotypic development is adaptive). If such counter-evidence exceeds what can be obtained over the remainder of ontogeny, plasticity terminates. If such counter-evidence exceeds what can be obtained within the next time period, but is feasible over the remainder of ontogeny, current experience can shape an organism's phenotype only at a later developmental stage (*lagged effects*) [20]. We refer to states—combinations of age, received cues and prior phenotypic development—as *critical* if the organism is insensitive to subsequent cues (i.e. phenotypic development has been set regardless of what cues follow).

When cues are poor (e.g. 0.55), the prior probability may affect plasticity. If the prior is 0.9, there is little reason to attend to cues; individuals commit to a phenotype early on and so an adoption does not result in phenotypic divergence. If the prior is 0.5, all the individuals have to go on are the poor cues. In this case, individuals retain plasticity through ontogeny, largely tracking noise, and so an adoption results in some phenotypic divergence. If the prior is 0.7, individuals either receive cues consistent with the prior and commit early on, or receive cues inconsistent with the prior and largely track noise. The combined result is a little phenotypic divergence.

(b) Optimal policies

We now turn to the optimal developmental policies (figure 2 and electronic supplementary material, figures A2.1–A2.8). Most policies are plastic (i.e. phenotypic decisions depend on estimates of the environmental state) at the onset of development. An exception occurs when cues are poor and the prior is 0.9; in this case, individuals largely ignore cues and rely on their prior.

Although most policies start out plastic, plasticity typically declines across ontogeny (i.e. fewer individuals switch phenotypic trajectories), and often drops to zero. As noted, this is why some doppelgängers do not phenotypically diverge despite receiving opposite cues. These critical states result from experience or from a combination of age and experience, but rarely from age alone (which occurs if plasticity drops to zero for all individuals at the same age, regardless of past experience). An age-dependent drop in plasticity occurs if the prior is 0.5, cues are weakly informative, and rewards are increasing (electronic supplementary material, figures A2.6–A2.8). Faced with a high probability of mismatch, organisms receive one cue and develop accordingly, trying to capture the rewards of expertise.

More commonly, plasticity depends on experience, and individual differences in plasticity result from stochasticity in cues [21,47]: some individuals obtain consistent cue sets and reach a critical state, becoming insensitive to cues early in ontogeny; others obtain heterogeneous cue sets and reach critical states only much later, if at all. In some cases, plasticity

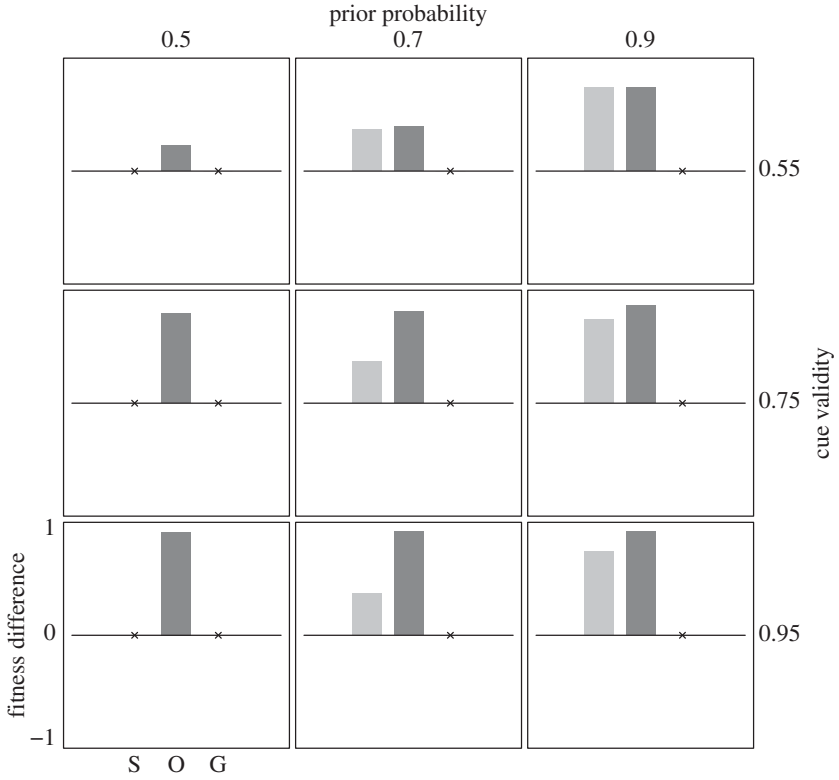

Figure 3. Fitness of mature phenotypes with linear rewards for correct phenotypic development and linear penalties for incorrect development. Each row of three panels depicts a cue validity (0.55, 0.75, 0.95) and each column of three panels a prior probability of E_1 (0.5, 0.7, 0.9). The height of each bar measures the expected fitness difference from baseline (an X denotes no difference). A fitness of 1 represents a perfect match with the environment. Negative fitness differences result from mismatch penalties exceeding benefits of proper calibration. The light-grey bar labelled 'S' denotes a pure specialist, the dark-grey bar 'O' the optimal policy, and the light-grey bar 'G' a pure generalist. See electronic supplementary material, appendix 3, for negative fitness differences and for non-zero fitness differences for 'G'.

depends on a combination of age and experience. As ontogeny proceeds, the scope for learning declines. As older individuals have fewer opportunities to obtain cues, they reach critical states with less certain estimates than younger individuals.

(c) Optimal policies versus non-plastic strategies

Plasticity may entail costs such as constructing and maintaining machinery for sensing, interpreting and acting on cues [31,48]. Natural selection should only favour plasticity if the benefits outweigh such costs. And so, it is informative to compare optimal policies, which by definition achieve expected fitness equal to or higher than any other strategy, with simpler alternatives. We benchmark optimal policies against two non-plastic strategies (figure 3 and electronic supplementary material, figures A3.1–A3.8): a 'pure specialist', which specializes maximally towards the environmental state with the higher prior probability (picking at random if the prior is 0.5), and a 'pure generalist', which specializes halfway towards each phenotypic target.

For pure specialists, only the prior affects fitness; for pure generalists, only the reward and penalty functions affect fitness (see electronic supplementary material, figures A3.1–A3.8). Both of these non-plastic strategies are unaffected by the cue validity, as cues do not shape their phenotypes. The prior affects the fitness of pure specialists as it determines what fraction of individuals specializes correctly, but not the fitness of pure generalists as they specialize to an equal extent (halfway) for each environmental state. The fitness of the optimal policy depends primarily on the cue validity: if cues are moderately or strongly informative, the optimal policy attains high fitness; if cues are weakly informative, the optimal policy's fitness increases with the prior.

The optimal policy performs well and strongly outperforms either non-plastic strategy when the prior is close to 0.5 and cues are moderately or strongly informative. In these cases, the optimal policy benefits maximally from the information provided by cues, outperforming pure specialists who develop 100% mismatched phenotypes 50% of the time, and outperforming pure generalists who develop 50% mismatched phenotypes 100% of the time.

(d) Phenotypic distributions

Different optimal policies generate different distributions of mature phenotypes (figure 4 and electronic supplementary material, figures A4.1–A4.8). If cues are highly informative, individuals rely heavily on them and most develop highly specialized phenotypes. In this case, ontogeny depends little on the prior (i.e. the data swamp the prior). In contrast, if cues are moderately informative, distributions are more dispersed: stochasticity in cues leads some individuals (those who obtain consistent cue sets) to develop highly specialized phenotypes; others (those who obtain heterogenous cue sets) either become generalists, which are adapted to some extent to both environmental states, or wait for some time, resulting in lower levels of specialization [21,47].

Despite receiving a steady stream of cues throughout ontogeny, the optimal policies may produce a substantial amount of phenotype–environment mismatch (electronic supplementary material, figures A5.1–A5.9). Mismatch is lower if cues are more informative. With poor cues and a prior of 0.7 or 0.9, individuals mostly rely on the prior and ignore cues. As a result, a fraction of the population adapts to the incorrect environment, becoming severely mismatched. With poor cues and a prior close to 0.5, the reward and penalty functions interact, giving

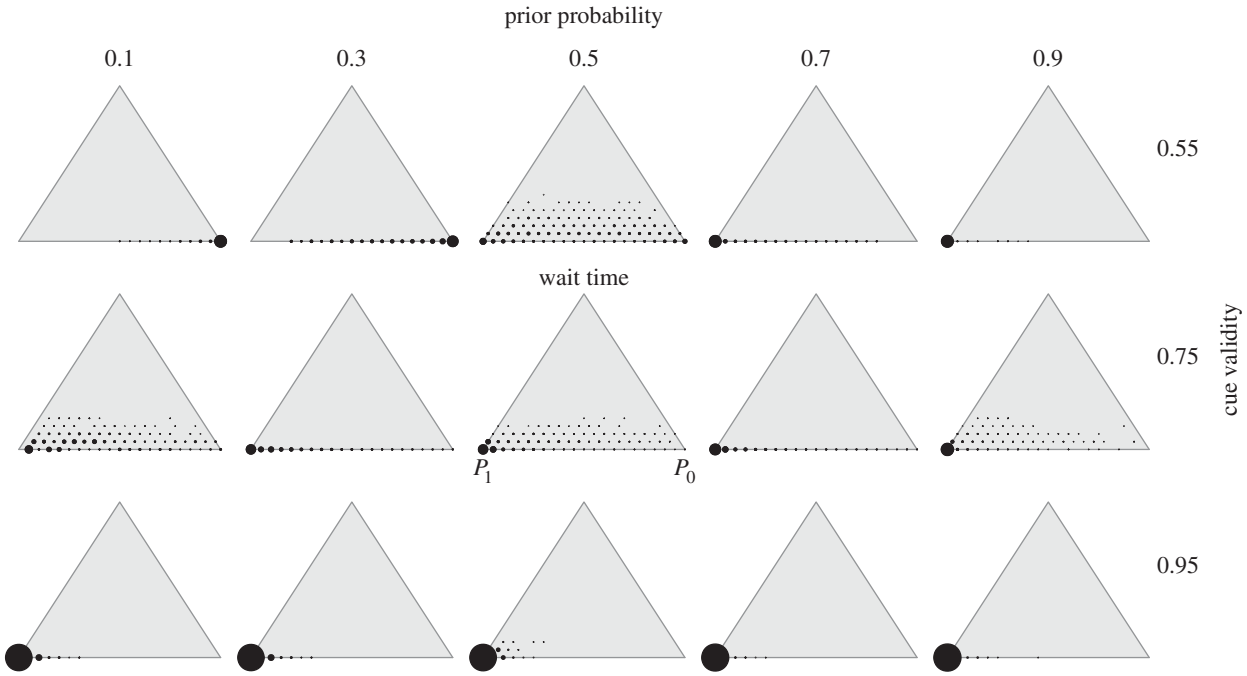

Figure 4. Distributions of mature phenotypes with linear rewards for correct phenotypic development and linear penalties for incorrect development. Each row of five panels depicts a cue validity (0.55, 0.75, 0.95) and each column of three panels a prior probability of E_1 (0.1, 0.3, 0.5, 0.7, 0.9). Each triangle plots a distribution of mature phenotypes, obtained by simulating 10 000 individuals following the optimal policy, developing in E_1. Phenotypes are defined by three numbers: time periods an organism waits, specializes towards P_1, and specializes towards P_0. The left vertex of a triangle represents complete specialization towards P_1; the right vertex complete specialization towards P_0; the top vertex waiting throughout development. The interior represents a mix of all three decisions; an edge a mix of two. A circle's area is proportional to the number of individuals of that phenotype.

rise either to individuals that track noise and develop generalist phenotypes (which are partially mismatched) or spend a fraction of ontogeny waiting before building their phenotypes.

4. Discussion

To understand the form–function fit in biology, we need models that capture both developmental and evolutionary dynamics. When the state of the environment varies, this fit is often achieved through phenotypic plasticity. Though there is a large literature on the evolutionary forces resulting in the presence or absence of plasticity, there has been comparatively little research on the evolution of sensitive periods.

One reason could be that some scholars regard sensitive periods as by-products of developmental processes [8,16], rather than adaptations. This view may be correct in some cases, but it does not address immense variation in sensitive periods: (i) between species in the same trait (e.g. some bird species learn their songs exclusively in early life, others throughout their entire lifetime; [49]), (ii) between individuals of the same species (e.g. individuals may vary in the extent to which the same experience shapes their development; [41]), (iii) between traits within a single individual (e.g. cognitive systems may adjust more easily than emotional systems to a radically changed environment; [11]), and (iv) within traits across development [19]. Our model generates such variation.

(a) Main findings

We reflect on our main results. First, matching the empirical record [4–12], we find that sensitive periods result from experience or a combination of age and experience, but rarely from age alone. The empirical record shows that sensitive periods can depend on experience in at least three different ways. Sensitive periods may be prolonged if: (i) organisms are deprived of any relevant cues [8,16,18]; (ii) features of cues are gradually changing [18,50]; or (iii) perceptual systems provide unstable inputs to the brain, because they are still developing or disrupted [16]. All three of these processes share a lack of reliable information available to the organism. Bateson & Martin [18] note: 'processes that bring the sensitive period to an end are related to the gathering of crucial information and, except in extreme cases, do not shut down until that information has been gathered' (p. 162).

Second, stochasticity in cues results in individual differences in the duration of sensitive periods; more consistent cue sets result in earlier loss of plasticity [21,47]. In our modelling framework, the noisiness of cues affects the 'gathering of crucial information' and prolongs plasticity. This is consistent, for example, with a study on auditory development in rats [51]. Pups exposed to a constant stream of white noise delay auditory specialization well beyond the critical period observed for rats reared in standard laboratory conditions.

Third, experience does not always result in immediate phenotypic consequence, but may manifest later in development. In our model, the environmental state remains stable across ontogeny. Related models, which incorporate environmental change across ontogeny, generate such *lagged effects* [20,26] for the same reason: having developed a certain posterior based on previous experience, individuals need substantial counter-evidence to revise their estimate (to the point where switching phenotypic development is adaptive).

Fourth, such counter-evidence may exceed what can be obtained over the remainder of ontogeny, in which case individuals will perseverate along developmental trajectories, despite mounting evidence suggesting they may be specializing incorrectly. These critical periods result, in part, from imposing a fixed end to ontogeny. Future work could explore

an extension of our model in which ontogeny ends with some probability in each time period.

(b) Robustness of findings

We impose a fitness penalty for environment–phenotype mismatch. We assume that the fitness penalty associated with each increment of mismatch equals the fitness reward associated with each correct developmental increment. We have analysed but not presented different relative penalty magnitudes. These yield similar qualitative results, with less waiting as penalties go to zero, and more as they increase. With no mismatch penalty, there is no reason to forgo phenotypic specialization. With a penalty magnitude larger than the reward, if the penalty diminishes (i.e. initial mismatch steps result in large fitness costs), selection may favour organisms that wait their whole lives, refusing to specialize for fear of mismatch.

We assume no cost of plasticity, but compare the optimal policy to two different non-plastic strategies. We find that selection favours phenotypic plasticity when there is high uncertainty early in development about the environmental state (e.g. the prior is close to 0.5) and cues are moderately or strongly informative [19]. If plasticity were costly, selection may favour non-plastic strategies (i.e. ignoring cues) when cues are weakly informative. Similarly, when initial uncertainty is low, it might pay to bet on the most likely outcome (the prior) rather than invest in costly machinery to reduce uncertainty. However, the costs of plasticity may vary across ontogeny (e.g. due to changing trade-offs between body parts) [19]. Introducing such a variable cost could favour more age-dependent plasticity.

(c) Limitations and future directions

We discuss two limitations. First, we assume the environmental state remains stable across ontogeny. If environmental change or migration occur during the lifetime [20,33], organisms must simultaneously infer the environmental state and consider environmental change, a more complicated inferential task. If environmental states are uncorrelated, selection may favour non-plastic phenotypes (specialists if the prior is close to 1 and generalists if close to 0.5). If environmental states are correlated across time but moving targets, plasticity may be prolonged. Moreover, the rate of environmental change may itself be a variable that developing organisms estimate and adjust to.

Second, we assume that the trait under investigation develops independently of all other traits (i.e. no evolutionary constraints on shaping developmental mechanism). This assumption constitutes a 'developmental gambit' analogous to the 'phenotypic gambit' [52] commonly assumed in behavioural ecology.

We are just beginning to catalogue the range of variation in phenotypic plasticity across species, individuals and traits, and to understand the underlying developmental and physiological mechanisms [53]. We need evolutionary theory to complete this picture. Our model of incremental development is part of the larger programme of evolutionary optimality theory, which predicts what organisms *should* do, but not *how* they do it. The evo-mecho approach represents a next step in ultimate-level theorizing, offering an explanation of how and why specific mechanisms evolve in specific ecologies [54,55]. And, the evo-mecho approach may help explain the evolution of developmental mechanisms capable of generating novel phenotypes [56], mechanisms capable of modifying their own environment [57] and even mechanisms capable of shaping their evolutionary futures [58].

Authors' contributions. Both K.P. and W.F. contributed to all components of this study.

Competing interests. We declare we have no competing interests.

Funding. This research was supported by a Veni grant from the Netherlands Organization for Scientific Research (NWO) (016.155.195) awarded to W.F.

Acknowledgements. We thank Martin Daly, Sinead English, Judy Stamps and four anonymous reviewers for their thoughtful comments.

References

1. West-Eberhard MJ. 2003 *Developmental plasticity and evolution*. New York, NY: Oxford University Press.
2. DeWitt TJ, Scheiner SM (eds). 2004 *Phenotypic plasticity: functional and conceptual approaches*. New York, NY: Oxford University Press.
3. Schlichting CD, Pigliucci M. 1998 *Phenotypic evolution: a reaction norm perspective*. Sunderland, MA: Sinauer.
4. Fagiolini M, Jensen CL, Champagne FA. 2009 Epigenetic influences on brain development and plasticity. *Curr. Opin. Neurobiol.* **19**, 207–212. (doi:10.1016/j.conb.2009.05.009)
5. Faulk C, Dolinoy DC. 2011 Timing is everything: the when and how of environmentally induced changes in the epigenome of animals. *Epigenetics* **6**, 791–797. (doi:10.4161/epi.6.7.16209)
6. Fox SE, Levitt P, Nelson CAIII. 2010 How the timing and quality of early experiences influence the development of brain architecture. *Child Dev.* **81**, 28–40. (doi:10.1111/j.1467-8624.2009.01380.x)
7. Knudsen EI. 2004 Sensitive periods in the development of the brain and behavior. *J. Cogn. Neurosci.* **16**, 1412–1425. (doi:10.1162/0898929042304796)
8. Michel GF, Tyler AN. 2005 Critical period: a history of the transition from questions of when, to what, to how. *Dev. Psychobiol.* **46**, 156–162. (doi:10.1002/dev.20058)
9. Sullivan RM, Holman PJ. 2010 Transitions in sensitive period attachment learning in infancy: the role of corticosterone. *Neurosci. Biobehav. Rev.* **34**, 835–844. (doi:10.1016/j.neubiorev.2009.11.010)
10. Takesian AE, Hensch TK. 2013 Balancing plasticity/stability across brain development. *Prog. Brain Res.* **207**, 3–34. (doi:10.1016/B978-0-444-63327-9.00001-1)
11. Zeanah CH, Gunnar MR, McCall RB, Kreppner JM, Fox NA. 2011 VI. Sensitive periods. *Monogr. SRCD* **76**, 147–162. (doi:10.1111/j.1540-5834.2011.00631.x)
12. Zevin JD. 2012 A sensitive period for shibboleths: the long tail and changing goals of speech perception over the course of development. *Dev. Psychobiol.* **54**, 632–642. (doi:10.1002/dev.20611)
13. Bateson P, Horn G. 1994 Imprinting and recognition memory: a neural net model. *Anim. Behav.* **48**, 695–715. (doi:10.1006/anbe.1994.1289)
14. Bornstein MH. 1989 Sensitive periods in development: structural characteristics and causal interpretations. *Psychol. Bull.* **105**, 179–197. (doi:10.1037/0033-2909.105.2.179)
15. Colombo J. 1982 The critical period concept: research, methodology, and theoretical issues. *Psychol. Bull.* **91**, 260–275. (doi:10.1037/0033-2909.91.2.260)
16. Thomas MSC, Johnson MH. 2008 New advances in understanding sensitive periods in brain

development. *Curr. Dir. Psychol. Sci.* **17**, 1–5. (doi:10.1111/j.1467-8721.2008.00537.x)

17. Bateson PPG. 1979 How do sensitive periods arise and what are they for? *Anim. Behav.* **27**, 470–486. (doi:10.1016/0003-3472(79)90184-2)
18. Bateson P, Martin P. 1999 *Design for a life: how behaviour develops*. London, UK: Jonathan Cape.
19. Fawcett TW, Frankenhuis WE. 2015 Adaptive explanations for sensitive windows in development. *Front. Zool.* **12**, S3. (doi:10.1186/1742-9994-12-S1-S3)
20. Fischer B, Van Doorn GS, Dieckmann U, Taborsky B. 2014 The evolution of age-dependent plasticity. *Am. Nat.* **183**, 108–125. (doi:10.1086/674008)
21. Frankenhuis WE, Panchanathan K. 2011 Balancing sampling and specialization: an adaptationist model of incremental development. *Proc. R. Soc. B* **278**, 3558–3565. (doi:10.1098/rspb.2011.0055)
22. English S, Fawcett T, Higginson A, Trimmer P, Uller T. In press. Adaptive use of information during growth can explain long-term effects of early-life experiences. *Am. Nat.* (doi:10.1086/685644)
23. Hurford JR. 1991 The evolution of the critical period for language acquisition. *Cognition* **40**, 159–201. (doi:10.1016/0010-0277(91)90024-X)
24. Komarova NL, Nowak MA. 2001 Natural selection of the critical period for language acquisition. *Proc. R. Soc. Lond. B* **268**, 1189–1196. (doi:10.1098/rspb.2001.1629)
25. Stamps JA, Krishnan VV. 2014 Combining information from ancestors and personal experiences to predict individual differences in developmental trajectories. *Am. Nat.* **184**, 647–657. (doi:10.1086/678116)
26. Utz M, Jeschke JM, Loeschcke V, Gabriel W. 2014 Phenotypic plasticity with instantaneous but delayed switches. *J. Theor. Biol.* **340**, 60–72. (doi:10.1016/j.jtbi.2013.08.038)
27. Snell-Rood EC. 2012 Selective processes in development: implications for the costs and benefits of phenotypic plasticity. *Integr. Comp. Biol.* **52**, 31–42. (doi:10.1093/icb/ics067)
28. Levins R. 1968 *Evolution in changing environments*. Princeton, NJ: Princeton University Press.
29. Moran NA. 1992 The evolutionary maintenance of alternative phenotypes. *Am. Nat.* **139**, 971–989. (doi:10.1086/285369)
30. Jablonka E, Oborny B, Molnár I, Kisdi E, Hofbauer J, Czárán T. 1995 The adaptive advantage of phenotypic memory. *Phil. Trans. R. Soc. Lond. B* **350**, 133–141. (doi:10.1098/rstb.1995.0147)
31. DeWitt TJ, Sih A, Wilson DS. 1998 Costs and limits of plasticity. *Trends Ecol. Evol.* **13**, 77–81. (doi:10.1016/S0169-5347(97)01274-3)
32. Gluckman PD, Hanson MA, Spencer HG, Bateson P. 2005 Environmental influences during development and their later consequences for health and disease: implications for the interpretation of empirical studies. *Proc. R. Soc. B* **272**, 671–677. (doi:10.1098/rspb.2004.3001)
33. Nettle D, Frankenhuis WE, Rickard IJ. 2013 The evolution of predictive adaptive responses in human life history. *Proc. R. Soc. B* **280**, 20131343. (doi:10.1098/rspb.2013.1343)
34. Dall SRX, Giraldeau L-A, Olsson O, McNamara JM, Stephens DW. 2005 Information and its use by animals in evolutionary ecology. *Trends Ecol. Evol.* **20**, 187–193. (doi:10.1016/j.tree.2005.01.010)
35. Donaldson-Matasci MC, Bergstrom CT, Lachmann M. 2010 The fitness value of information. *Oikos* **119**, 219–230. (doi:10.1111/j.1600-0706.2009.17781.x)
36. McNamara JM, Dall SRX. 2010 Information is a fitness enhancing resource. *Oikos* **119**, 231–236. (doi:10.1111/j.1600-0706.2009.17509.x)
37. Agrawal AA, Laforsch C, Tollrian R. 1999 Transgenerational induction of defences in animals and plants. *Nature* **401**, 60–63. (doi:10.1038/43425)
38. Frankenhuis WE, Panchanathan K, Barrett HC. 2013 Bridging developmental systems theory and evolutionary psychology using dynamic optimization. *Dev. Sci.* **16**, 584–598. (doi:10.1111/desc.12053)
39. Houston AI, McNamara JM. 1999 *Models of adaptive behaviour: an approach based on state*. Cambridge, MA: Cambridge University Press.
40. Mangel M, Clark CW. 1988 *Dynamic modeling in behavioral ecology*. Princeton, NJ: Princeton University Press.
41. Frankenhuis W, Panchanathan K, Belsky J. 2015 A mathematical model of the evolution of individual differences in developmental plasticity arising through parental bet-hedging. *Dev. Sci.* (doi:10.1111/desc.12309)
42. McNamara JM, Green RF, Olssen O. 2006 Bayes' theorem and its applications in animal behavior. *Oikos* **112**, 243–251. (doi:10.1111/j.0030-1299.2006.14228)
43. McNamara J, Houston AI. 1980 The application of statistical decision theory to animal behaviour. *J. Theor. Biol.* **85**, 673–690. (doi:10.1016/0022-5193(80)90265-9)
44. Trimmer PC, Houston AI, Marshall JAR, Mendl MT, Paul ES, McNamara JM. 2011 Decision-making under uncertainty: biases and Bayesians. *Anim. Cogn.* **14**, 465–476. (doi:10.1007/s10071-011-0387-4)
45. Kooi RE, Brakefield PM. 1999 The critical period for wing pattern induction in the polyphenic tropical butterfly *Bicyclus anynana* (Satyrinae). *J. Insect Physiol.* **45**, 201–212. (doi:10.1016/S0022-1910(98)00093-6)
46. Rountree DB, Nijhout HF. 1995 Hormonal control of a seasonal polyphenism in *Precis coenia* (Lepidoptera: Nymphalidae). *J. Insect Physiol.* **41**, 987–992. (doi:10.1016/0022-1910(95)00046-W)
47. Frankenhuis WE, Panchanathan K. 2011 Individual differences in developmental plasticity may result from stochastic sampling. *Perspect. Psychol. Sci.* **6**, 336–347. (doi:10.1177/1745691611412602)
48. Auld JR, Agrawal AA, Relyea RA. 2010 Reevaluating the costs and limits of adaptive phenotypic plasticity. *Proc. R. Soc. B* **277**, 503–511. (doi:10.1098/rspb.2009.1355)
49. Beecher MD, Brenowitz EA. 2005 Functional aspects of song learning in songbirds. *Trends Ecol. Evol.* **20**, 143–149. (doi:10.1016/j.tree.2005.01.004)
50. Bolhuis JJ. 1991 Mechanisms of avian imprinting: a review. *Biol. Rev.* **66**, 303–345. (doi:10.1111/j.1469-185X.1991.tb01145.x)
51. Chang EF, Merzenich MM. 2003 Environmental noise retards auditory cortical development. *Science* **300**, 498–502. (doi:10.1126/science.1082163)
52. Fawcett TW, Hamblin S, Giraldeau L-A. 2013 Exposing the behavioral gambit: the evolution of learning and decision rules. *Behav. Ecol.* **24**, 2–11. (doi: 10.1093/beheco/ars085)
53. Nijhout HF. 2015 A developmental–physiological perspective on the development and evolution of phenotypic plasticity. In *Conceptual change in biology* (ed. AC Love), pp. 147–173. Dordrecht, The Netherlands: Springer.
54. McNamara JM, Houston AI. 2009 Integrating function and mechanism. *Trends Ecol. Evol.* **24**, 670–675. (doi: 10.1016/j.tree.2009.05.011)
55. Fawcett TW, Fallenstein B, Higginson AD, Houston AI, Mallpress DEW, Trimmer PC, McNamara JM. 2014 The evolution of decision rules in complex environments. *Trends Cogn. Sci.* **18**, 153–161. (doi:10.1016/j.tics.2013.12.012)
56. Frank SA. 1996 The design of natural and artificial adaptive systems. In *Adaptation* (eds MR Rose, GV Lauder), pp. 451–505. New York, NY: Academic Press.
57. Flynn EG, Laland KN, Kendal RL, Kendal JR. 2013 Developmental niche construction. *Dev. Sci.* **16**, 296–313. (doi:10.1111/desc.12030)
58. Oyama S, Griffiths PE, Gray RD (eds) 2001 *Cycles of contingency: developmental systems and evolution*. Cambridge, MA: MIT Press.