

# UC Davis

## UC Davis Previously Published Works

### Title

Combining information from ancestors and personal experiences to predict individual differences in developmental trajectories.

### Permalink

<https://escholarship.org/uc/item/8w92z386>

### Journal

The American naturalist, 184(5)

### ISSN

0003-0147

### Authors

Stamps, Judy A  
Krishnan, VV

### Publication Date

2014-11-01

### DOI

10.1086/678116

Peer reviewed



CHICAGO JOURNALS



The University of Chicago

---

Combining Information from Ancestors and Personal Experiences to Predict Individual Differences in Developmental Trajectories

Author(s): Judy A. Stamps and V. V. Krishnan,

Source: *The American Naturalist*, (-Not available-), p. 000

Published by: [The University of Chicago Press](#) for [The American Society of Naturalists](#)

Stable URL: <http://www.jstor.org/stable/10.1086/678116>

Accessed: 02/10/2014 11:49

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



*The University of Chicago Press, The American Society of Naturalists, The University of Chicago* are collaborating with JSTOR to digitize, preserve and extend access to *The American Naturalist*.

<http://www.jstor.org>

# Combining Information from Ancestors and Personal Experiences to Predict Individual Differences in Developmental Trajectories

Judy A. Stamps<sup>1,\*</sup> and V. V. Krishnan<sup>2</sup>

1. Evolution and Ecology, University of California, Davis, California; 2. School of Engineering, San Francisco State University, San Francisco, California

Submitted November 13, 2013; Accepted June 6, 2014; Electronically published October 1, 2014

Online enhancements: appendixes.

**ABSTRACT:** A persistent question in biology is how information from ancestors combines with personal experiences over the lifetime to affect the developmental trajectories of phenotypic traits. We address this question by modeling individual differences in behavioral developmental trajectories on the basis of two assumptions: (1) differences among individuals in the behavior expressed at birth or hatching are based on information from their ancestors (via genes, epigenes, and prenatal maternal effects), and (2) information from ancestors is combined with information from personal experiences over ontogeny via Bayesian updating. The model predicts relationships between the means and the variability of the behavior expressed by neonates and the subsequent developmental trajectories of their behavior when every individual is reared under the same environmental conditions. Several predictions of the model are supported by data from previous studies of behavioral development, for example, that the temporal stability of personality will increase with age and that the intercepts and slopes of developmental trajectories for boldness will be negatively correlated across individuals or genotypes when subjects are raised in safe environments. We describe how other specific predictions of the model can be used to test the hypothesis that information from ancestors and information from personal experiences are combined via nonadditive, Bayesian-like processes.

**Keywords:** innate, predispositions, developmental systems, differential consistency, intraindividual variability, IIV, Bayesian updating, personality, boldness, repeatability.

## Introduction

A long-standing assumption in biology is that information that shapes the development of behavior can come from a variety of different sources. Information can come from an individual's distant ancestors (e.g., via genes; Leimar et al. 2006; Shea 2007) or from its immediate ancestors

(e.g., via maternal effects [Uller 2008] or inherited epigenetic markers [Bonduriansky and Day 2009; Shea et al. 2011]). Of course, information can also come from personal experiences that animals have over the course of their lifetimes, for example, via learning (Shettleworth 2010) or the many other processes by which information from the environment can affect development (reviews in West-Eberhard 2003; Gluckman et al. 2005*b*; Bateson and Gluckman 2011).

What is not clear is how information from an individual's ancestors and information from its personal experiences combine with each other over ontogeny to affect the expression and development of phenotypic traits. The lack of attention to this question has left empiricists who study individual differences in developmental patterns with observations in search of explanations. For instance, investigators who use longitudinal protocols to study personality in humans and animals have found that personality is less temporally stable (less differentially consistent) early in life than later in life (squid: Sinn et al. 2008; humans: Roberts and DelVecchio 2000; Caspi et al. 2005; dogs: Fratkin et al. 2013; fish: Edenbrow and Croft 2011), but the reasons for this pattern are currently unclear. Similarly, researchers have detected significant positive or negative correlations across individuals or genotypes between trait values and the developmental plasticity of those traits (Auld et al. 2010; Mathot et al. 2012), but again, the reasons for these patterns are obscure.

To date there have been two suggestions about ways that information from ancestors might combine with information from personal experiences to affect developmental trajectories. Classical quantitative genetics begins with the assumption that the effects of genes and experiential factors on phenotypic traits are additive (Falconer and Mackay 1996), and this assumption has been incorporated into models showing that selection can favor the evolution

\* Corresponding author; e-mail: jastamps@ucdavis.edu.

of developmental mechanisms that combine information from genetic cues and information from environmental cues to predict the conditions that are likely to occur during an individual's lifetime (Leimar 2005, 2009; Leimar et al. 2006). Other authors have suggested that information from ancestors, evolutionary history, or genes might be combined with information from personal experiences via nonadditive, Bayesian-like processes (e.g., Dall et al. 2005; McNamara et al. 2006; Pierre and Green 2008; Schmidt et al. 2010). This idea follows from the widespread success of Bayesian approaches to model learning (McNamara and Houston 1980; Dall et al. 2005; Courville et al. 2006; McLinn and Stephens 2006; Trimmer et al. 2011) and from the fact that, in principle, Bayesian updating is the best way to combine information from different sources to estimate the state of the world (Lange and Dukas 2009).

Bayesian models for combining information from ancestors with information from personal experience sound promising, but at present, such models are still in their infancy. Frankenhuis and Panchanathan (2011*a*, 2011*b*) made an important first step in addressing this issue, by using Bayesian approaches to demonstrate how stochastic variation in sampling might lead to individual differences in developmental trajectories, even if every individual starts with the same information from their ancestors (i.e., every individual has the same prior distribution). Their model provides a plausible explanation for intragenotypic variability, situations in which virtually isogenic animals, raised under the same set of conditions, express different behavior (Freund et al. 2013; Stamps et al. 2013). However, it does not address the broader question of how developmental trajectories would be affected if individuals began with different prior distributions, based on information from their ancestors.

Two problems, one minor and one major, must be addressed to investigate how individual differences in information from ancestors might combine with information from personal experiences to affect trait development. First, any general model of development must account for the fact that individuals or genotypes that express the same mean trait values at a given age can differ in their developmental plasticity, that is, the extent to which their trait values change after exposure to the same external stimuli (e.g., Cohen et al. 2008; Auld et al. 2010; Dingemans and Wolf 2013). For instance, individuals who express the same level of antipredator behavior at a given age can express very different levels of antipredator behavior after exposure to cues from a predator (Bell and Sih 2007). In Bayesian models, this issue can be easily addressed using prior or posterior distributions that differ with respect to both their means and their variances (see "Model Description" and apps. A, E; apps. A–E available online).

The second, more challenging problem is designing a

testable model. A major sticking point is finding a way to estimate the prior distributions of individuals before they have been exposed to a particular type of personal experience. Empiricists using Bayesian models to study learning typically sidestep this problem, either by assuming that every individual has the same, noninformative prior distribution at the beginning of the experiment (e.g., Holyoak and Cheng 2011) or by thoroughly pretraining their subjects to ensure that they all have the same prior distribution before measuring their behavior (e.g., McLinn and Stephens 2006; Biernaskie et al. 2009). Here we suggest one possible solution to this problem: assume that the behavior expressed by naive individuals, before they have had any relevant personal experience, is based on information from their ancestors. Behavioral biologists studying development routinely make this assumption (e.g., Adret 2004; Bremner 2011; Westerman et al. 2012; Waters and Burgardt 2013). By extension, if individuals not only express a given type of behavior soon after birth or hatching but also continue to express that behavior as they grow and develop, it seems reasonable to assume that personal experiences during the juvenile period might affect the developmental trajectory of that behavior. Building on these assumptions, we show how one can estimate the mean and the variance of the prior distributions of neonates on the basis of the mean and the variability of the behavior they express soon after birth or hatching. Then, standard Bayesian approaches can be used to predict the behavioral developmental trajectories of different individuals as a function of the behavior they expressed as neonates and to predict how those developmental trajectories would change as a function of the conditions in which those individuals were raised. This approach can be used to generate predictions that can be readily tested by empiricists, using protocols already available to study individual differences in behavioral traits.

We illustrate this approach by analyzing the developmental trajectories for behavior patterns related to "boldness," a personality trait that has been studied in a wide range of animals, including humans (Fox et al. 2005; Réale et al. 2007; Conrad et al. 2011). One advantage of focusing on boldness is that in many species, individuals exhibit different levels of boldness soon after birth or hatching, before they have been exposed to personal experiences that might provide information about the level of danger in their environment (e.g., see Edenbrow and Croft 2011; Sussman and Ha 2011). In addition, it is possible to estimate, at least at a qualitative level, information relevant to danger that might be conveyed to developing individuals by cues they perceive in their environment. For instance, it seems reasonable to assume that repeated exposure to cues from predators over ontogeny might indicate that the environment was relatively dangerous.

Conversely, the absence of cues from predators, aggressive conspecifics, or any other potential dangers might also convey information about the current environment (Sih 1992; Welton et al. 2003; Stamps et al. 2009). In particular, it seems reasonable to assume that the continued absence of any cues indicative of danger over ontogeny would indicate to developing animals that the current environment is relatively safe.

The goals of this study were to explore the implications of assuming that information from ancestors and information from personal experiences are combined by Bayesian updating and, in particular, to design a simple model based on this assumption whose predictions could be readily tested by empiricists studying individual differences in the development of behavioral and other traits. We first describe the model and its predictions and then consider conditions that have to be satisfied to test those predictions. We show that several of those predictions are already supported by empirical data and discuss others that could be readily tested in a controlled laboratory setting. Finally, we conclude with suggestions for future research on the question of how information from ancestors and experience combines across ontogeny to affect the developmental trajectories of behavioral and other traits.

### Model Description

Since excellent introductions to the use of Bayesian models for studying animal behavior are available elsewhere (see McNamara and Houston 1980; Hilborn and Mangel 1997; McNamara et al. 2006; Trimmer et al. 2011), here we summarize the concepts most relevant to this study. Bayesian models include four basic components: prior distributions, posterior distributions, likelihood functions, and response functions. Informally, a prior distribution specifies an individual's beliefs about a biologically relevant variable (e.g., the state of danger) before it has a given experience (e.g., being chased by a predator), and a posterior distribution specifies that individual's beliefs about that same variable after it has had that experience. The likelihood function for a particular type of experience specifies the probability that that experience would occur, given each possible state of the variable; the response function links beliefs to behavior, by specifying the relationship between an individual's current belief (based on its prior or posterior distribution) and the behavior it expresses based on that belief. The posterior distribution after one experience becomes the prior distribution for the next experience, which is why Bayesian approaches are useful for modeling development, where it is typical for a given individual to have a series of experiences over ontogeny, each of which may provide additional information about a state of the world (Frankenhuis and Panchanathan

2011a, 2011b). When Bayesian models are empirically tested, most authors do not assume that individuals behave in a strictly Bayesian fashion. Instead, it is typically assumed that humans or animals use rules of thumb or other cognitive processes that yield reasonable approximations of the estimates of patterns of behavior that would be generated by Bayesian updating (McNamara and Houston 1980; Gigerenzer and Todd 1999; McNamara et al. 2006; Trimmer et al. 2011; Bowers and Davis 2012). For additional background on the Bayesian approach used in this study, see appendix A.

Here, we consider how animals might combine information from ancestors and information from personal experiences to estimate the state of a variable we call "danger," where danger indicates the risk of injury, harm, or death. We assume that in nature, the state of danger can take on a value ranging from a minimum value in the safest possible environment to a maximum value in the most dangerous possible environment. For computational ease, we assume that there are 100 possible states of danger, ranging in increments of 0.01 from a minimum value of 0 (no danger) to a maximum value of 1.0 (the maximum level of danger in the natural environment), and we describe prior distributions and posterior distributions that specify the probability that the level of danger falls within each of the 100 mutually exclusive possible states (see app. B).

One of the difficulties in describing prior distributions is that, in theory, these distributions can take on a variety of different shapes. For our model we use the beta distribution to generate a wide range of biologically reasonable prior distributions. Beta distributions use two parameters,  $\alpha$  and  $\beta$ , to generate probabilities whose values continuously vary in the range between 0 and 1. We focus on sets of  $\alpha$  and  $\beta$  values that generate monotonically increasing, monotonically decreasing, unimodal (hump-shaped), and uniform distributions. We do not consider U-shaped distributions ( $\alpha < 1$  and  $\beta < 1$ ), in which the extreme states of 0 and 1 are both more likely to occur than any of the intermediate states.

This study asks how differences among agents in the means and the variances of their prior distributions would affect the behavior of those agents after they all had the same experience (same likelihood function). We focused on prior distributions with mean values ranging from 0.1 to 0.9. For each of these mean values, we set the lowest variance at 0.001 and determined the beta distribution that generated this variance. For each mean value, we also determined the beta distribution with the highest possible variance, under the constraint noted above (beta distributions with  $\alpha \geq 1$  and/or  $\beta \geq 1$ ).

Beta distributions were also used to model biologically reasonable likelihood functions for a given variable. Like-

likelihood functions can vary with respect to their “reliability” (McLinn and Stephens 2006) or “validity” (Frankenhuis and Panchanathan 2011*a*), terms that indicate the probability of a given experience being associated with the different possible states of the variable. We assume that any personal experiences or cues that generate likelihood functions are readily perceived by every individual, so that we can ignore sampling costs or individual differences in estimates of the state due to sampling error (see Frankenhuis and Panchanathan 2011*b*). In the main text, we illustrate the basic concepts using a simple right-biased likelihood function (i.e., one that indicates that the experience is more strongly associated with states with high values than with states with low or moderate values), but we consider a range of other possible likelihood functions in appendix C.

Response functions link prior or posterior distributions with the behavior that an agent expresses based on those distributions. When prior and posterior distributions are continuously distributed, the mean (expected value) of those distributions is generally considered to provide the best estimator of the state (Trimmer et al. 2011), so we assume that the mean level of behavior expressed by an agent at any given time is directly related to the mean of its prior or posterior distribution at that point in time. In the text, we illustrate our main points by using linear response functions, in which the mean level of behavior is linearly related to the mean of the prior or posterior distribution. However, we have also examined response functions with other shapes (sigmoid, exponential, and asymptotic; see app. D).

An important innovation is our assumption that the variability of the behavior of an agent at a given time is positively related to the variance of its prior or posterior distribution at that point in time. At the individual level, the stochastic, short-term variability in the behavior expressed in a given context is termed “intraindividual variability” (IIV; Nesselroade 1991; Stamps et al. 2012; Wang et al. 2012; Biro and Adriaenssens 2013). At the genotypic level, the stochastic variability in the behavior of isogenic individuals reared under the same conditions and then tested at the same age in the same context has been called “intragenotypic variability” (Stamps et al. 2013). Recent empirical studies have revealed significant differences across individuals in IIV (Stamps et al. 2012; Biro and Adriaenssens 2013; Briffa et al. 2013) and significant differences across genotypes in intragenotypic variability (Kain et al. 2012; Stamps et al. 2013), indicating that it is practical to detect differences among agents in either type of behavioral variability, if such differences exist. Importantly, there is empirical support for the assumption that IIV is related to the variance of prior distributions. Recent studies of animals and humans indicate that subjects can

be trained to estimate the variance of prior distributions and that when this is done, the IIV of their neural activity and/or their behavioral responses is positively related to the variance of those prior distributions (Daw et al. 2005; Berniker et al. 2010; Funamizu et al. 2012).

In order to model how information from ancestors affects behavioral developmental trajectories, we assume that agents with all of the prior distributions described above (mean values from 0.1 to 0.9, variances from 0.001 to the maximum possible variance for each mean) exist in the same population, that the mean behavior of each neonate is directly related to the mean of its prior distribution, and that the variability of its behavior is directly related to the variance of its prior distribution. Each agent is then exposed to the same experience (i.e., an experience with a given likelihood function). We combine the prior distribution with the likelihood function, using Bayesian updating to obtain a posterior distribution, as described in appendix B, and we assume that the mean and variance of the behavior each agent expresses after exposure to the experience are directly related to the mean and variance of its posterior distribution. This process allows us to estimate how a single experience would affect the means and variances of the posterior distributions of agents with a wide range of prior distributions (see app. E). We then repeat this process and plot the output of the response function for each of the prior or posterior distributions to predict changes over ontogeny in the expected behavior and the variability of behavior of each agent.

For each agent, we can compute a developmental trajectory, which describes the expected (mean) level of behavior expressed by that agent at a series of ages. The expected behavior at birth or hatching (i.e., at age 0) is given by the intercept of an agent’s developmental trajectory, and the change in behavior after one or more experiences is indicated by the shape or slope of its developmental trajectory. We can also determine how the behavioral variability of an agent changes over ontogeny, by plotting the intraindividual or intragenotypic variability of behavior for each agent at age 0 and at each successive age. Analyses were run for the full range of prior distributions indicated above (see app. E), but for purposes of illustration in the text we focus on 15 agents, each with a different prior distribution (five means of 0.1, 0.3, 0.5, 0.7 or 0.9 and three variances of 0.001, 0.02, and the maximum possible variance for the given mean). Together, these 15 distributions span the range of prior distributions that are possible under the assumptions of our model. We assume that each individual is exposed four times to the same experience (same likelihood function), and then we graph the effects of that experience on its developmental trajectory and behavioral variability over ontogeny.

## Results and Discussion

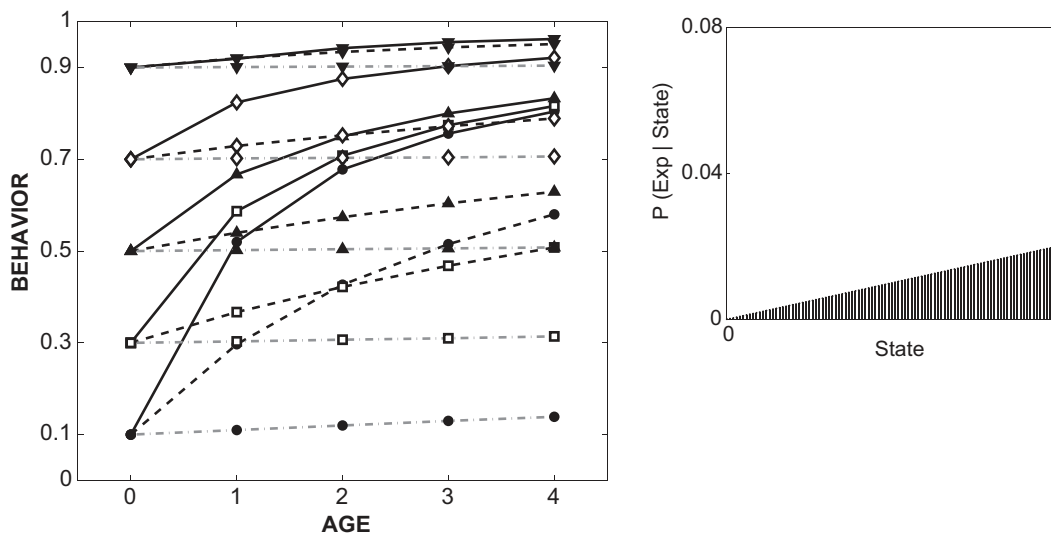
### How Different Prior Distributions Affect Developmental Trajectories

We can now consider how developmental trajectories of agents with different prior distributions would change if they were all repeatedly exposed to a right-biased likelihood function (fig. 1). One obvious result is that most developmental trajectories are nonlinear: for those agents whose behavior changes over ontogeny, the rate of change in behavior is higher early in ontogeny than later in ontogeny. In addition, the shapes of the trajectories vary systematically across agents. In the current example, in which the likelihood function is right biased, ontogenetic changes in behavior very early in life (age 0–1) are most pronounced for agents whose prior distributions have low means and high variances, and they are least pronounced for agents whose prior distributions have very high means (e.g., 0.9) or very low variances (0.001). In addition, when the slopes of developmental trajectories are computed across the entire period (from age 0 to age 4), the slopes and intercepts of the trajectories are negatively related to one another across the 15 agents. Figure 1 also shows that differences among agents in prior distributions can have long-lasting effects on their behavior later in life. That is, despite repeated exposure to an experience with a moderately reliable likelihood function, agents with very low

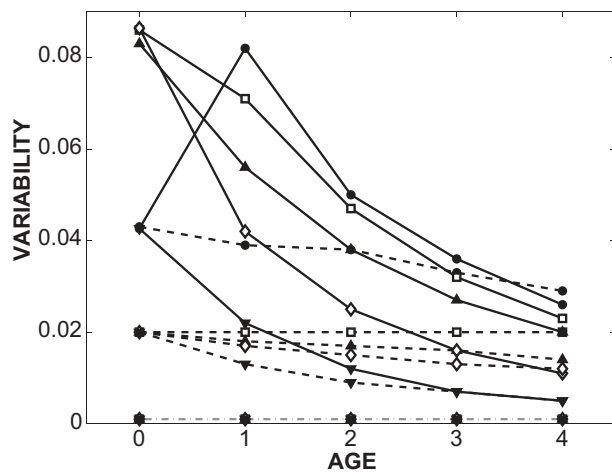
mean scores for behavior at age 0 still have lower scores at age 4 than agents who had high scores at age 0.

Finally, figure 1 indicates that the temporal consistency of individual differences in behavior is lower early in ontogeny (age 0–1) than later in ontogeny (age 3–4). Differential consistency (also called “broad-sense repeatability”) indicates the extent to which individual differences in behavior are maintained over a specified period of time (Hayes and Jenkins 1997; Caspi and Roberts 2001; Stamps and Groothuis 2010). The differential consistency of behavior over a given period (e.g., from age 0 to age 1) can be estimated by computing the slope of each agent’s developmental trajectory over that period and then computing the variance, across agents, in their slopes. Differential consistency is negatively related to the variance in slopes across agents, so that the highest possible value of differential consistency occurs when every agent has the same slope. In the example illustrated in figure 1, the slopes of the developmental trajectories vary more across agents early in ontogeny (age 0–1) than they do later in ontogeny (age 3–4), indicating that differential consistency increases with age.

We next consider how the variability of behavior (IIV for individuals or intragenotypic variability for genotypes) would change over ontogeny if agents with different prior distributions were repeatedly exposed to the same, right-biased experience (fig. 2). When agents have prior distri-



**Figure 1:** *Left*, developmental trajectories (based on the expected behavior at five ages) for 15 hypothetical agents (individuals or genotypes), after repeated exposure to an experience with a moderately reliable, right-biased likelihood function (shown on *right*), assuming a linear response function. The means of their prior distributions (at age 0) are indicated by the following symbols: circles for 0.1, squares for 0.3, upward triangles for 0.5, diamonds for 0.7, and downward triangles for 0.9. The variances of their prior distributions are indicated by lines: dot-dashed gray for 0.001, dashed black for 0.02, and solid black for the maximum variance for the given mean. *Right*, the likelihood function indicates the probability of the experience, given the state ( $P(\text{Exp}|\text{State})$ ), for each of the 100 possible states, ranging from 0 to 1.



**Figure 2:** Change over ontogeny in the behavioral variability (intraindividual variability for individuals, intragenotypic variability for genotypes) of 15 hypothetical agents after repeated exposure to an experience with a moderately reliable, right-biased likelihood function (fig. 1, *right*), assuming a linear response function. Symbols and lines indicate prior means and prior variances as in figure 1. All of the agents with prior distributions with very low behavioral variability (0.001) are indicated by the same dot-dashed gray line because for all of them, behavioral variability is maintained at levels near 0 throughout their lives.

butions with high variance (indicated by solid lines), the variability of behavior is high at age 0, after which it usually declines over ontogeny. A notable exception is the agent with a low prior mean and a high prior variance (indicated by circles and solid black lines in fig. 2), whose behavioral variability increases from age 0 to age 1 and then gradually declines later in ontogeny. This result follows from a general principle of Bayesian updating, namely, that information that sharply conflicts with an individual's previous estimate of the state of the world is likely to decrease, rather than increase, its level of certainty about its estimate of the state (see app. E; fig. E2; figs. A1, A2, C1–C7, D1–D4, E1, E2 available online). For agents whose prior distributions have intermediate variance (0.02, indicated by dashed lines), behavioral variability declines modestly, if at all, over ontogeny. Finally, in agents with prior distributions with low variance (0.001, indicated by dot-dashed lines), variability is very low at age 0 and remains at the same low level for the rest of ontogeny. Across all of the agents, the model predicts that both average variability and interindividual differences in variability would be higher early in ontogeny (from age 0 to age 1) than later in ontogeny (from age 3 to age 4).

At a qualitative level, most of these patterns are robust to changes in likelihood and response functions (see apps. C, D). Across all of the likelihood functions and response

functions tested, our model indicates that if animals are repeatedly exposed to the same experience, that is, the same likelihood function, then (1) the rate of change of behavior is higher early in ontogeny than later in ontogeny, (2) the temporal consistency of differences among agents in mean levels of behavior (differential consistency) is lower earlier in ontogeny than later in ontogeny, (3) the rate of change in behavior over ontogeny is highest for agents whose mean level of behavior when naive differs most from the behavior encouraged by the likelihood function, (4) differences among agents in their prior distributions continue to affect the behavior expressed by those individuals later in life, (5) on average, behavioral variability declines over ontogeny, and (6) agents with low behavioral variability at birth or hatching change their behavior less over ontogeny than do agents with high behavioral variability at birth or hatching.

However, one important pattern that does change as a function of likelihood functions is the relationship, across agents, between the absolute value (i.e., the magnitude) of the slope,  $S$ , and the intercept,  $I$ , of their developmental trajectories. Right-biased likelihood functions encourage negative relationships between  $I$  and  $S$  (e.g., figs. 1, C1–C4). However, if the same set of agents were repeatedly exposed to an experience with a left-biased likelihood function, the model predicts a positive relationship between  $I$  and  $S$  (e.g., fig. C6), while repeated exposure to a unimodal (hump-shaped) likelihood function is expected to lead to a nonlinear relationship, across agents, between  $I$  and  $S$  (e.g., fig. C7). Thus, for the same set of agents, correlations between the intercepts and the magnitude of the slopes of their behavioral developmental trajectories are predicted to vary as a function of the likelihood function for the conditions in which they were raised.

#### *Testing Predictions of a Bayesian Model of Development*

**Assumptions.** An important consideration in testing any Bayesian model of behavioral development is choosing the right type of personal experience. This is not a trivial problem, because many types of experiences, especially those that occur early in life, constrain the development of behavioral and other traits by restricting the resources available for somatic growth and development (Monaghan 2008; Bateson and Gluckman 2011; Nettle et al. 2013). Examples include periods of food deprivation or bouts of infection soon after birth or hatching. Such experiences might provide information about food levels or risk of parasitism later in life, but they also have immediate effects on the level or quality of resources that can be allocated to the development of the physiological and morphological systems that generate behavior (Gluckman et al. 2005a; Monaghan 2008; Devevey et al. 2010). Situations in which



a given experience provides information about conditions later in life but also has immediate effects on resource allocation are more complicated to model and study (Monaghan 2008; Nettle et al. 2013). Hence, empirical tests of Bayesian models of development should focus on personal experiences that provide organisms with information about the state of the world but do not directly affect the resources that are available to support somatic growth and development.

A second important implicit assumption is that the state of the world remains constant over the period when behavior is measured. In the laboratory, it may be possible to ensure that environmental conditions do not change through ontogeny. However, this assumption must also be approximately valid under natural conditions for the study species. If in nature the state of the world is highly variable over ontogeny, individuals might weigh recent information more heavily than old information when estimating the current state (Mangel 1990). Alternately, an individual's own internal state might provide it with information about predictable changes in the state of the world over ontogeny. For instance, if small juveniles are more vulnerable to predators than large juveniles (e.g., Sogard 1997), an individual's current size might provide it with indirect, relatively reliable information about its current risk of predation. Similarly, if juveniles are less likely than adults to be attacked by conspecific adults (delBarco-Trillo et al. 2011; Templeton et al. 2012), an individual's maturational state might provide it with information about its risk of being attacked. In support of these ideas, there is evidence that developmental trajectories for boldness can change when individuals undergo major transitions in morphological or physiological state (e.g., during maturation [Sinn et al. 2008] or across metamorphosis [Hedrick and Kortet 2012]), even if those animals are maintained in captivity under constant conditions. Thus, the predictions of the current model are most likely to apply when the state of variables such as danger does not vary predictably over ontogeny in the natural habitat.

A final caveat is that all of the subjects in an experiment should have equal access to cues that provide information about the state of the world. Under natural conditions, this may not be the case, for example, bolder individuals might be more likely to sample the environment, and hence gain more information about it, than shy ones (see Mathot et al. 2012). Practically speaking, this means using cues that can be readily perceived by every subject and do not require active sampling.

*Predictions.* A general prediction from our model is that when different individuals or genotypes develop under the same set of conditions, differential consistency (broad-sense repeatability) will increase with age. Our model

therefore provides a simple general explanation for the ontogenetic increases in the temporal stability of personality that have been reported in several taxa (see "Introduction").

More important, our model predicts specific patterns of developmental trajectories when agents who express different levels of behavior when naive are raised under specific sets of environmental conditions. For instance, we suggest that an extended period with no cues from predators, aggressive conspecifics, or other dangers would, under natural conditions, indicate that an individual was living in a relatively safe environment. Conveniently, many researchers studying the development of behavioral traits, including boldness, raise and maintain their subjects under standard laboratory conditions. In part because of animal welfare concerns and in part to reduce variability in behavior, standard laboratory conditions typically lack any cues or stimuli from predators, aggressive conspecifics, or other biotic or abiotic cues that might be associated with risk or danger in nature. Of course, tests used to assess boldness necessarily involve exposing subjects to stimuli they perceive to be at least modestly dangerous. However, it is possible to assay boldness using stimuli that do not reliably indicate high levels of danger, to conduct tests infrequently, and to maintain the experimental subjects under benign conditions between each test (e.g., Edenbrow and Croft 2011). Thus, assuming that an individual's level of boldness would be negatively related to its estimate of the level of danger in the environment (see app. A), our model predicts that repeated exposure to experiences suggesting (with intermediate reliability) that the environment was safe would generate developmental trajectories for boldness similar to those illustrated in figure 1.

Although studies of individual or genotypic differences in the developmental trajectories of boldness are still quite rare, we found two that appear to satisfy the criteria outlined above. Recently, Sussman and Ha (2011) reported on a study of developmental trajectories for boldness in pigtailed macaques (*Macaca nemestrina*). Their subjects were maintained in individual cages, with no exposure to predators or dangerous conspecifics from birth through the first 10 months of life, and the monkeys exhibited significant differences in boldness as neonates. As predicted by our model, the rate of change in boldness was highest in infants and declined at older ages: the juveniles' developmental trajectories had the shapes predicted by the model (best fitted using growth equations in which behavior scores approached asymptotes at older ages). In addition, as predicted by our model, there was a significant negative relationship ( $r = -0.52$ ,  $P < .001$ ) across 152 individuals between their intercepts and slopes: infants who were initially shy changed more (became much bolder) over the 10-month period, while infants who were

initially bold did not change as much (remained relatively bold) over the same period of time (A. Sussman, personal communication).

In a second example, Edenbrow and Croft (2011) studied the ontogeny of boldness in mangrove killifish (*Kryptolebias marmoratus*), taking advantage of the fact that this is one of a handful of vertebrates that produce clones. They raised fish from 20 genotypes in isolation from hatching, with no cues from predators, including conspecifics (cannibalism has been reported in this species; see Taylor 2012). The fish began to mature at about 90 days, so we focus here on the results of tests of boldness for juveniles at days 2, 30, and 61 posthatch. Levels of boldness at day 2 varied significantly across the genotypes (M. Edenbrow, personal communication). Subsequently, boldness increased nonlinearly as a function of age, with higher rates of change for young juveniles than for older juveniles and significant differences among the genotypes in the slopes of their developmental trajectories for boldness (Edenbrow and Croft 2011). In this case, there was a nonsignificant negative relationship across the 20 genotypes between boldness scores at day 2 and the rate of increase in boldness from day 2 to day 61 ( $r = -0.371$ ,  $P = .107$ ; M. Edenbrow, personal communication).

Several predictions of our model have not yet been tested. One is that the developmental trajectories of the same agents would vary in specific ways if they were raised under conditions with different likelihood functions. For instance, if the set of killifish clones described in the previous paragraph were repeatedly exposed during the juvenile period to cues from predators, we predict that their developmental trajectories for boldness would be similar to those illustrated in figure C6. That is, individuals who initially were very bold would become shyer, while individuals who were initially shy would remain shy. In contrast, if the same clones were raised with cues indicating an intermediate level of danger, we predict that their trajectories would be similar to those illustrated in figure C7: very bold individuals would become shyer, very shy individuals would become bolder, and intermediately bold individuals would maintain their behavior. That is, our model predicts that relationships between the intercepts and the magnitude of the slopes of developmental trajectories for the same set of agents can be negative or positive or have other shapes, depending on the likelihood function. Thus, it provides a possible explanation for observed variation across studies and across species in the extent and direction of correlations between trait values and the developmental plasticity of those traits (Auld et al. 2010; Mathot et al. 2012).

Several novel, nonintuitive predictions of the model follow from our assumption that behavioral variability at birth or hatching is related to the variance of the prior

distribution, that is, the extent to which neonates are “certain” about their prior estimate of the state of the world (see app. A). Here, the major caveat is that the behavior in question should be fully functional in neonates, since the intraindividual variability of some types of behavior may change over ontogeny as a result of maturational changes in neurological or morphological systems (see the ontogenetic changes in the IIV of reaction times in humans; Tamnes et al. 2012). However, if we can assume that the variability of the behavior expressed by a neonate reflects that individual’s uncertainty about its estimate of the state of the world, several predictions are possible. For instance, our model indicates that for agents who express the same mean level of behavior when naive, those with high levels of behavioral variability as neonates would tend to change their mean behavior more over ontogeny than those with low levels of behavioral variability at the same age. The model also predicts that when agents express intermediate levels of behavioral variability as neonates, behavioral variability would initially increase with age for agents whose mean behavior as neonates was very different from the level of behavior encouraged by the experience but would decline with age for agents whose mean behavior as neonates was closer to the level of behavior encouraged by the experience.

### General Discussion and Conclusions

This study provides a simple but potentially powerful way to begin investigating a long-standing question in biology, namely, how information from ancestors combines with information from personal experiences to generate individual or genotypic differences in developmental patterns. In the absence of attention to this question, it has been difficult for empiricists to interpret patterns they have observed in their data and difficult for theoreticians to incorporate reasonable assumptions about this process into their models.

We illustrated this approach with a simple scenario, in which individual differences in estimates of one variable (here, danger) affect the development of one type of behavior (here, behavior related to boldness). However, assuming that the basic conditions are met (e.g., behavior is expressed by neonates and then continues to be expressed by juveniles), the same approach could be extended to generate testable predictions involving other continuously distributed variables (e.g., population density), other behavioral traits (e.g., aggressiveness), or even physiological traits that vary among neonates, continue to be expressed over ontogeny, and change in response to cues from the environment (e.g., developmental trajectories for stress responsiveness and the hypothalamic-pituitary-adrenal axis; Hostinar et al. 2014; Koch et al.

2014). In addition, of course, individual differences in prior distributions might affect the developmental trajectories of trait syndromes involving combinations of behavioral, physiological, and morphological traits (see Kasumovic 2013). However, since morphological traits often respond more slowly and less reversibly to changes in environmental cues than do behavioral traits (Gabriel et al. 2005; Hossie and Murray 2012), more complicated models involving lagged or constrained responses might be required to predict the developmental trajectories of morphological traits under different sets of environmental conditions, as a function of differences among agents in their prior distributions at birth or hatching.

More generally, we suggest that a better understanding of how information from ancestors combines with information from personal experiences over ontogeny will be necessary to convert the study of individual or genotypic differences in developmental patterns from a descriptive to a predictive science. Despite many years of discussion about ways that genes, maternal effects, personal experiences, and other factors might interact over ontogeny to affect the development of behavioral and other traits (see Oyama 2000; Bateson and Gluckman 2011), empirical studies of individual or genotypic differences in developmental trajectories or of relationships between the intercepts and slopes of developmental trajectories are still primarily descriptive, because there is little theory to drive them. We hope that this article will encourage others to consider how information from ancestors might combine with information from personal experiences over ontogeny to affect developmental trajectories. For instance, it might be useful to construct models based on different assumptions about rules for combining information from ancestors and information from experience (e.g., models based on additive rules) and then ask which models do a better job of predicting individual or genotypic differences in developmental trajectories. Once we have a better idea of how information from ancestors and information from personal experiences combine over ontogeny to affect developmental trajectories, it will be possible to incorporate reasonable assumptions about this process into future studies of the proximate and ultimate factors responsible for the interindividual or intergenotypic variation in developmental trajectories that is so often observed for behavioral and other traits.

#### Acknowledgments

We are very grateful to A. Bell, P. Biro, W. Frankenhuys, and two anonymous reviewers for extensive comments on previous versions of the manuscript and to M. Edenbrow and A. Sussman for kindly providing us with unpublished

results from their studies of the development of boldness in killifish and pigtailed macaques, respectively.

#### Literature Cited

- Adret, P. 2004. In search of the song template. Pages 303–324 in H. P. Zeigler, and P. Marler, eds. Behavioral neurobiology of birdsong. Annals of the New York Academy of Sciences no. 1016.
- Auld, J. R., A. A. Agrawal, and R. A. Relyea. 2010. Re-evaluating the costs and limits of adaptive phenotypic plasticity. *Proceedings of the Royal Society B: Biological Sciences* 277:503–511.
- Bateson, P., and P. Gluckman. 2011. Plasticity, robustness, development and evolution. Cambridge University Press, New York.
- Bell, A. M., and A. Sih. 2007. Exposure to predation generates personality in threespined sticklebacks (*Gasterosteus aculeatus*). *Ecology Letters* 10:828–834.
- Berniker, M., M. Voss, and K. Kording. 2010. Learning priors for Bayesian computations in the nervous system. *PLoS ONE* 5:e12686. doi:10.1371/journal.pone.0012686.
- Biernaskie, J. M., S. C. Walker, and R. J. Gegear. 2009. Bumblebees learn to forage like Bayesians. *American Naturalist* 174:413–423.
- Biro, P. A., and B. Adriaenssens. 2013. Predictability as a personality trait: consistent differences in intraindividual behavioral variation. *American Naturalist* 182:621–629.
- Bonduriansky, R., and T. Day. 2009. Nongenetic inheritance and its evolutionary implications. *Annual Review of Ecology, Evolution, and Systematics* 40:103–125.
- Bowers, J. S., and C. J. Davis. 2012. Bayesian just-so stories in psychology and neuroscience. *Psychological Bulletin* 138:389–414.
- Bremner, J. G. 2011. Four themes from 20 years of research on infant perception and cognition. *Infant and Child Development* 20:137–147.
- Briffa, M., D. Bridger, and P. A. Biro. 2013. How does temperature affect behaviour? multilevel analysis of plasticity, personality and predictability in hermit crabs. *Animal Behaviour* 86:47–54.
- Caspi, A., and B. W. Roberts. 2001. Personality development across the life course: the argument for change and continuity. *Psychological Inquiry* 12:49–66.
- Caspi, A., B. W. Roberts, and R. L. Shiner. 2005. Personality development: stability and change. *Annual Review of Psychology* 56:453–484.
- Cohen, H., A. B. Geva, M. A. Matar, J. Zohar, and Z. Kaplan. 2008. Post-traumatic stress behavioural responses in inbred mouse strains: can genetic predisposition explain phenotypic vulnerability? *International Journal of Neuropsychopharmacology* 11:331–349.
- Conrad, J. L., K. L. Weinersmith, T. Brodin, J. B. Saltz, and A. Sih. 2011. Behavioural syndromes in fishes: a review with implications for ecology and fisheries management. *Journal of Fish Biology* 78:395–435.
- Courville, A. C., N. D. Daw, and D. S. Touretzky. 2006. Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences* 10:294–300.
- Dall, S. R. X., L. A. Giraldeau, O. Olsson, J. M. McNamara, and D. W. Stephens. 2005. Information and its use by animals in evolutionary ecology. *Trends in Ecology and Evolution* 20:187–193.
- Daw, N. D., Y. Niv, and P. Dayan. 2005. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience* 8:1704–1711.

- delBarco-Trillo, J., M. E. McPhee, and R. E. Johnston. 2011. Syrian hamster males below an age threshold do not elicit aggression from unfamiliar adult males. *Aggressive Behavior* 37:91–97.
- Devevey, G., P. Bize, S. Fournier, E. Person, and P. Christe. 2010. Testing the predictive adaptive response in a host-parasite system. *Functional Ecology* 24:178–185.
- Dingemanse, N. J., and M. Wolf. 2013. Between-individual differences in behavioural plasticity within populations: causes and consequences. *Animal Behaviour* 85:1031–1039.
- Edenbrow, M., and D. P. Croft. 2011. Behavioural types and life history strategies during ontogeny in the mangrove killifish, *Kryptolebias marmoratus*. *Animal Behaviour* 82:731–741.
- Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. Benjamin-Cummings, San Francisco.
- Fox, N. A., H. A. Henderson, P. J. Marshall, K. E. Nichols, and M. M. Ghera. 2005. Behavioral inhibition: linking biology and behavior within a developmental framework. *Annual Review of Psychology* 56:235–262.
- Frankenhuis, W. E., and K. Panchanathan. 2011a. Balancing sampling and specialization: an adaptationist model of incremental development. *Proceedings of the Royal Society B: Biological Sciences* 278:3558–3565.
- . 2011b. Individual differences in developmental plasticity may result from stochastic sampling. *Perspectives on Psychological Science* 6:336–347.
- Fratkin, J. L., D. L. Sinn, E. A. Patall, and S. D. Gosling. 2013. Personality consistency in dogs: a meta-analysis. *PLoS ONE* 8: e54907. doi:10.1371/journal.pone.0054907.
- Freund, J., A. M. Brandmaier, L. Lewejohann, I. Kirste, M. Kritzler, A. Krüger, N. Sachser, U. Lindenberger, and G. Kempermann. 2013. Emergence of individuality in genetically identical mice. *Science* 340:756–759.
- Funamizu, A., M. Ito, K. Doya, R. Kanzaki, and H. Takahashi. 2012. Uncertainty in action-value estimation affects both action choice and learning rate of the choice behaviors of rats. *European Journal of Neuroscience* 35:1180–1189.
- Gabriel, W., B. Luttbeg, A. Sih, and R. Tollrian. 2005. Environmental tolerance, heterogeneity, and the evolution of reversible plastic responses. *American Naturalist* 166:339–353.
- Gigerenzer, G., and P. M. Todd. 1999. Simple heuristics that make us smart. Oxford University Press, New York.
- Gluckman, P. D., M. A. Hanson, and H. G. Spencer. 2005a. Predictive adaptive responses and human evolution. *Trends in Ecology and Evolution* 20:527–533.
- Gluckman, P. D., M. A. Hanson, H. G. Spencer, and P. Bateson. 2005b. Environmental influences during development and their later consequences for health and disease: implications for the interpretation of empirical studies. *Proceedings of the Royal Society B: Biological Sciences* 272:671–677.
- Hayes, J. P., and S. H. Jenkins. 1997. Individual variation in mammals. *Journal of Mammalogy* 78:274–293.
- Hedrick, A. V., and R. Kortet. 2012. Sex differences in the repeatability of boldness over metamorphosis. *Behavioral Ecology and Sociobiology* 66:407–412.
- Hilborn, R., and M. Mangel. 1997. The ecological detective: confronting models with data. Princeton University Press, Princeton, NJ.
- Holyoak, K. J., and P. W. Cheng. 2011. Causal learning and inference as a rational process: the new synthesis. *Annual Review of Psychology* 62:135–163.
- Hossie, T. J., and D. L. Murray. 2012. Assessing behavioural and morphological responses of frog tadpoles to temporal variability in predation risk. *Journal of Zoology* 288:275–282.
- Hostinar, C. E., R. M. Sullivan, and M. R. Gunnar. 2014. Psychological mechanisms underlying the social buffering of the hypothalamic-pituitary-adrenocortical axis: a review of animal models and human studies across development. *Psychological Bulletin* 140:256–282.
- Kain, J. S., C. Stokes, and B. L. de Bivort. 2012. Phototactic personality in fruit flies and its suppression by serotonin and *white*. *Proceedings of the National Academy of Sciences of the USA* 109: 19834–19839.
- Kasumovic, M. M. 2013. The multidimensional consequences of the juvenile environment: towards an integrative view of the adult phenotype. *Animal Behaviour* 85:1049–1059.
- Koch, H., K. McCormack, M. M. Sanchez, and D. Maestriperi. 2014. The development of the hypothalamic-pituitary-adrenal axis in rhesus monkeys: effects of age, sex, and early experience. *Developmental Psychobiology* 56:86–95.
- Lange, A., and R. Dukas. 2009. Bayesian approximations and extensions: optimal decisions for small brains and possibly big ones too. *Journal of Theoretical Biology* 259:503–516.
- Leimar, O. 2005. The evolution of phenotypic polymorphism: randomized strategies versus evolutionary branching. *American Naturalist* 165:669–681.
- . 2009. Environmental and genetic cues in the evolution of phenotypic polymorphism. *Evolutionary Ecology* 23:125–135.
- Leimar, O., P. Hammerstein, and T. J. M. Van Dooren. 2006. A new perspective on developmental plasticity and the principles of adaptive morph determination. *American Naturalist* 167:367–376.
- Mangel, M. 1990. Dynamic information in uncertain and changing worlds. *Journal of Theoretical Biology* 146:317–332.
- Mathot, K. J., J. Wright, B. Kempnaers, and N. J. Dingemanse. 2012. Adaptive strategies for managing uncertainty may explain personality-related differences in behavioural plasticity. *Oikos* 121:1009–1020.
- McLinn, C. M., and D. W. Stephens. 2006. What makes information valuable: signal reliability and environmental uncertainty. *Animal Behaviour* 71:1119–1129.
- McNamara, J., and A. Houston. 1980. The application of statistical decision theory to animal behavior. *Journal of Theoretical Biology* 85:673–690.
- McNamara, J. M., R. F. Green, and O. Olsson. 2006. Bayes' theorem and its applications in animal behaviour. *Oikos* 112:243–251.
- Monaghan, P. 2008. Early growth conditions, phenotypic development and environmental change. *Philosophical Transactions of the Royal Society B: Biological Sciences* 363:1635–1645.
- Nesselroade, J. R. 1991. The warp and woof of the developmental fabric. Pages 213–240 *in* R. Downs, L. Liben, and D. Palermo, eds. *Visions of development, the environment, and aesthetics: the legacy of Joachim F. Wohlwill*. Erlbaum, Hillsdale, NJ.
- Nettle, D., W. E. Frankenhuis, and I. J. Rickard. 2013. The evolution of predictive adaptive responses in human life history. *Proceedings of the Royal Society B: Biological Sciences* 280:20131343. doi: 10.1098/rspb.2013.1343.
- Oyama, S. 2000. The ontogeny of information: developmental systems and evolution. Duke University Press, Durham, NC.
- Pierre, J.-S., and R. F. Green. 2008. A Bayesian approach to optimal foraging in parasitoids. Pages 357–383 *in* É. Wajnberg, C. Bern-

- stein, and J. van Alphen, eds. Behavioral ecology of insect parasitoids. Blackwell, Oxford.
- Réale, D., S. M. Reader, D. Sol, P. T. McDougall, and N. J. Dingemanse. 2007. Integrating animal temperament within ecology and evolution. *Biological Reviews* 82:291–318.
- Roberts, B. W., and W. F. DelVecchio. 2000. The rank-order consistency of personality traits from childhood to old age: a quantitative review of longitudinal studies. *Psychological Bulletin* 126:3–25.
- Schmidt, K. A., S. R. X. Dall, and J. A. van Gils. 2010. The ecology of information: an overview on the ecological significance of making informed decisions. *Oikos* 119:304–316.
- Shea, N. 2007. Representation in the genome and in other inheritance systems. *Biology and Philosophy* 22:313–331.
- Shea, N., I. Pen, and T. Uller. 2011. Three epigenetic information channels and their different roles in evolution. *Journal of Evolutionary Biology* 24:1178–1187.
- Shettleworth, S. 2010. *Cognition, evolution and behavior*. Oxford University Press, New York.
- Sih, A. 1992. Prey uncertainty and the balancing of antipredator and feeding needs. *American Naturalist* 139:1052–1069.
- Sinn, D. L., S. D. Gosling, and N. A. Moltischniowskyj. 2008. Development of shy/bold behaviour in squid: context-specific phenotypes associated with developmental plasticity. *Animal Behaviour* 75:433–442.
- Sogard, S. M. 1997. Size-selective mortality in the juvenile stage of teleost fishes: a review. *Bulletin of Marine Science* 60:1129–1157.
- Stamps, J., and T. G. G. Groothuis. 2010. The development of animal personality: relevance, concepts and perspectives. *Biological Reviews* 85:301–325.
- Stamps, J., B. Luttbegg, and V. V. Krishnan. 2009. Effects of survival on the attractiveness of cues to natal dispersers. *American Naturalist* 173:41–46.
- Stamps, J. A., M. Briffa, and P. A. Biro. 2012. Unpredictable animals: individual differences in intraindividual variability (IIV). *Animal Behaviour* 83:1325–1334.
- Stamps, J. A., J. B. Saltz, and V. V. Krishnan. 2013. Genotypic differences in behavioural entropy: unpredictable genotypes are composed of unpredictable individuals. *Animal Behaviour* 86:641–649.
- Sussman, A., and J. Ha. 2011. Developmental and cross-situational stability in infant pigtailed macaque temperament. *Developmental Psychology* 47:781–791.
- Tamnes, C. K., A. M. Fjell, L. T. Westlye, Y. Østby, and K. B. Walhovd. 2012. Becoming consistent: developmental reductions in intraindividual variability in reaction time are related to white matter integrity. *Journal of Neuroscience* 32:972–982.
- Taylor, D. S. 2012. Twenty-four years in the mud: what have we learned about the natural history and ecology of the mangrove rivulus, *Kryptolebias marmoratus*? *Integrative and Comparative Biology* 52:724–736.
- Templeton, C. N., S. E. Campbell, and M. D. Beecher. 2012. Territorial song sparrows tolerate juveniles during the early song-learning phase. *Behavioral Ecology* 23:916–923.
- Trimmer, P. C., A. I. Houston, J. A. R. Marshall, M. T. Mendl, E. S. Paul, and J. M. McNamara. 2011. Decision-making under uncertainty: biases and Bayesians. *Animal Cognition* 14:465–476.
- Uller, T. 2008. Developmental plasticity and the evolution of parental effects. *Trends in Ecology and Evolution* 23:432–438.
- Wang, L. J., E. Hamaker, and C. S. Bergeman. 2012. Investigating inter-individual differences in short-term intra-individual variability. *Psychological Methods* 17:567–581.
- Waters, R. M., and G. M. Burghardt. 2013. Prey availability influences the ontogeny and timing of chemoreception-based prey shifting in the striped crayfish snake, *Regina alleni*. *Journal of Comparative Psychology* 127:49–55.
- Welton, N. J., J. M. McNamara, and A. I. Houston. 2003. Assessing predation risk: optimal behaviour and rules of thumb. *Theoretical Population Biology* 64:417–430.
- West-Eberhard, M. J. 2003. *Developmental plasticity and evolution*. Oxford University Press, New York.
- Westerman, E. L., A. Hodgins-Davis, A. Dinwiddie, and A. Monteiro. 2012. Biased learning affects mate choice in a butterfly. *Proceedings of the National Academy of Sciences of the USA* 109:10948–10953.

Associate Editor: Marc Mangel  
 Editor: Judith L. Bronstein

## **Appendix A from J. A. Stamps and V. V. Krishnan, “Combining Information from Ancestors and Personal Experiences to Predict Individual Differences in Developmental Trajectories” (Am. Nat., vol. 184, no. 5, p. 000)**

### **Effects of the Means and Variances of Prior Distributions on Bayesian Updating**

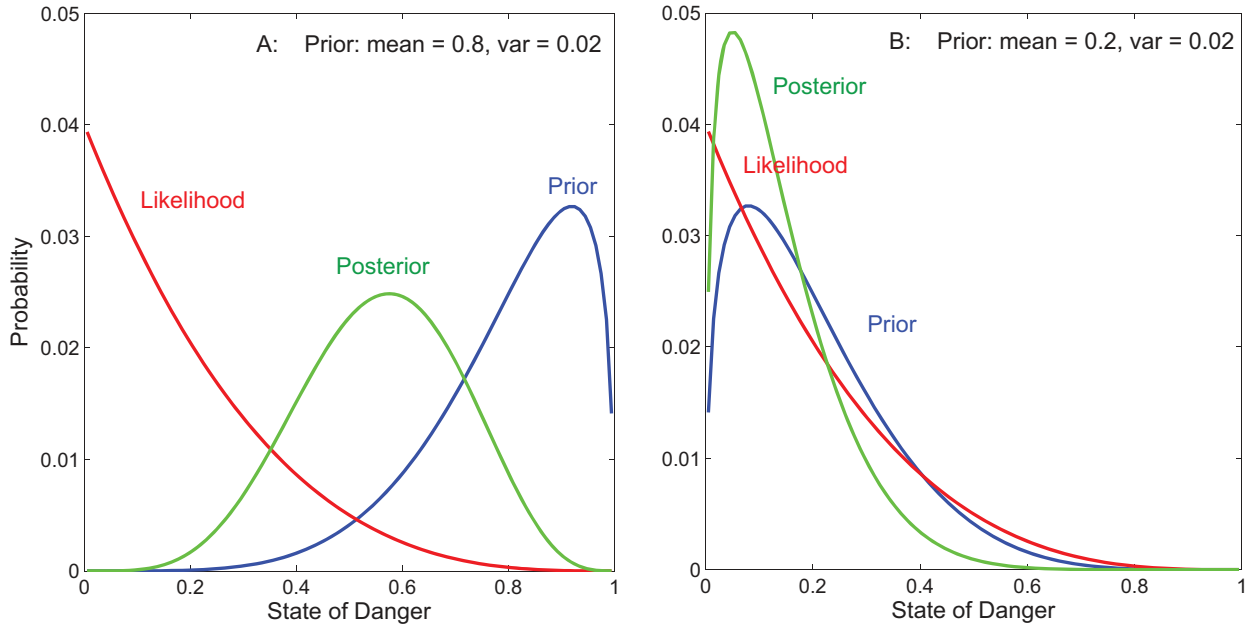
Here, we show how differences between individuals in the means of their prior distributions affect Bayesian updating. One individual (fig. A1A) has a prior distribution with a high mean, indicating that at birth or hatching, this individual estimates that the level of danger is more likely to be high than it is to be low or moderate. This individual is then maintained for a period (e.g., a month) under “safe” conditions, with no exposure to cues from predators, aggressive conspecifics, or any other potential dangers. As is indicated by the shape of the likelihood function, this type of experience is more likely to occur when the level of danger is low than when it is moderate to high. When this individual’s prior distribution is combined with this likelihood function, via Bayesian updating, it yields the posterior distribution shown in figure A1A. That is, after the experience, this individual revises downward its belief about the level of danger. Finally, assuming that “boldness” is negatively related to the mean of the prior or the posterior distribution for danger, we would expect this individual to be bolder after the experience (based on the mean of its posterior distribution) than it was when it was naive (based on the mean of its prior distribution).

A second individual (fig. A1B) has a prior distribution with a low mean, indicating that when it is naive, this individual estimates that the level of danger is more likely to be low than it is to be moderate to high. This individual is then exposed to the same experience (same likelihood function) as the individual in figure A1A. However, in this case, the estimate of the state of danger provided by this individual’s prior distribution is very similar to the estimate of the state of danger provided by the experience. As a result, its posterior distribution is very similar to its prior distribution. By extension, we would expect this individual’s boldness score after the experience to be similar to its score when it was naive.

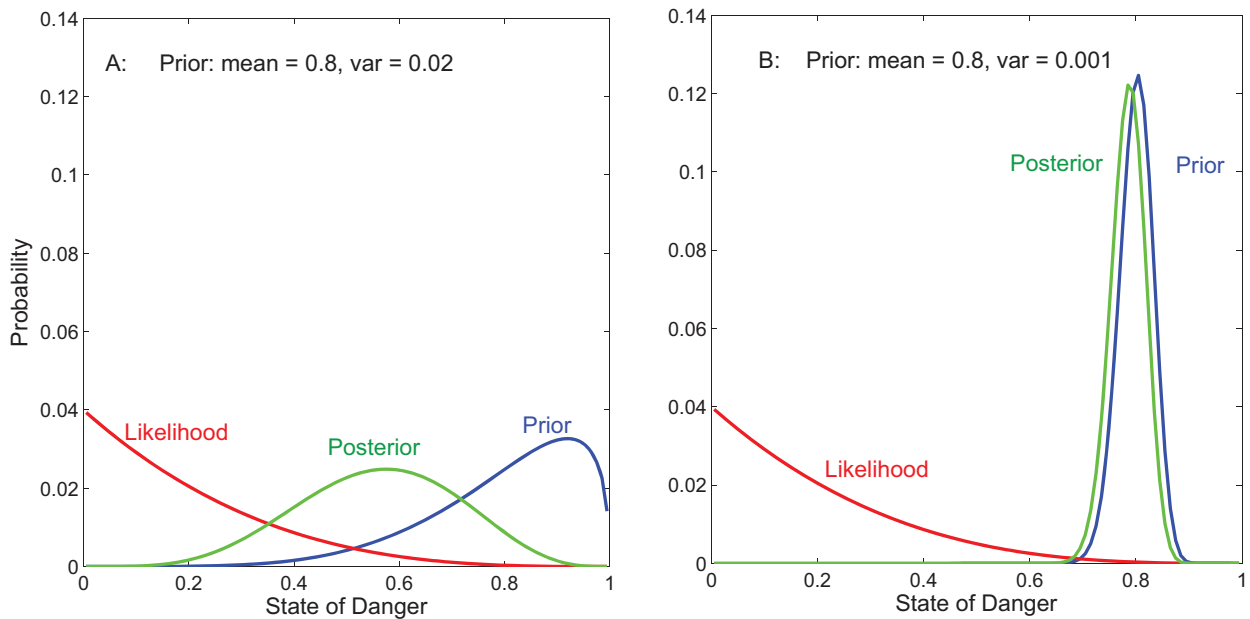
This example illustrates a very general, very basic feature of Bayesian updating, namely, that the effects of the same experience on estimates of the state of the world depend on the discrepancy between the prior distribution and the likelihood function (Courville et al. 2006). One can intuitively see that if a naive individual believes that the world is a safe place, an extended period of time with no cues indicative of danger simply confirms its initial belief and hence has little or no effect on its belief that the world is safe. On the other hand, if a naive individual believes that the world is dangerous, an extended period of time with no cues indicative of danger is a “surprise,” so this experience is more likely to change its estimate of danger.

A second important point is that the effect of a potentially informative experience on an individual’s estimate of the state of the world also depends on the variance of its prior distribution (fig. A2). Consider a situation in which two individuals both have prior distributions with the same mean value (mean = 0.8). That is, when naive, both of them estimate that the state of danger is relatively high. However, the variance of the first individual’s prior distribution (fig. A2A) is much higher than the variance of the second individual’s prior distribution (fig. A2B). Both individuals are then exposed to experience indicating that the level of danger is moderately low. In the case of the first individual, this experience leads to a reduction in its estimate of the level of danger, that is, a posterior distribution shifted to the left of its prior distribution (fig. A2A). However, in the case of the second individual, the same experience has little effect on its estimate of danger; its posterior distribution is very similar to its prior distribution (fig. A2B). Thus, although both individuals would be expected to express the same low level of boldness when naive, after the same experience, the first individual’s level of boldness would increase, but the second individual’s level of boldness would not change.

In this case, the intuitive explanation is that the variance of an individual’s prior distribution indicates the degree of confidence an individual has in its initial belief about the state of the world. If a naive individual vaguely suspects that the world might be dangerous, experience indicating that it is actually safe should alter its estimate of the state of the danger, and hence its behavior. However, if a naive individual firmly believes that the world is dangerous, that same experience should have little or no effect on either its belief or its behavior.



**Figure A1:** Effect of the mean of the prior distribution on Bayesian updating. Two individuals are both exposed to the same experience, with the likelihood function indicated in red. The first individual has a prior distribution with a high mean (*A*); the second individual has a prior distribution with a low mean (*B*). When the likelihood function and the prior distribution contradict each other (*A*), the posterior distribution is displaced from the prior distribution. In contrast, when the likelihood function and the prior distribution are concordant (*B*), the posterior distribution is very similar to the prior distribution.



**Figure A2:** Effect of the variance of the prior distribution on Bayesian updating. Two individuals are both exposed to the same experience, with the likelihood function indicated in red. Both individuals have prior distributions with the same mean, but the individual in *A* has a prior distribution with a high variance, whereas the individual in *B* has a prior distribution with a low variance. When variance of the prior is high (*A*), the posterior distribution is displaced from the prior distribution. However, when the variance of the prior is low (*B*), the posterior distribution is virtually the same as the prior distribution.

# Appendix B from J. A. Stamps and V. V. Krishnan, “Combining Information from Ancestors and Personal Experiences to Predict Individual Differences in Developmental Trajectories” (Am. Nat., vol. 184, no. 5, p. 000)

## Estimating Posterior Probabilities Using Bayesian Estimation

Our analysis is based on the assumption that an individual’s behavior is influenced by its estimate of the value of a variable  $A$  (e.g., danger) and that these values can vary continuously between a minimum and a maximum. However, for the purposes of computing the posterior probability, it is easier to discretize the variable into a number of small, mutually exclusive states. We assume that the minimum and maximum values of the variable are  $A_{\min} = 0$  and  $A_{\max} = 1$ . Thus, when we discretize the variable into  $n = 100$  equally spaced states between its minimum and maximum values, the  $i$ th state will be bounded by  $(0.01(i - 1), 0.01i)$ . If  $A_1, A_2, \dots, A_n$  are the  $n$  individual states, then the probability of the variable being in state  $A_i$  is described by  $P(A_i)$ . Also, since  $A_1, A_2, \dots, A_n$  represent all possible states of the variable, the sum of their probabilities will add up to 1, that is,

$$P(A_1) + P(A_2) + \dots + P(A_n) = 1.$$

Next,  $P(B)$  is the probability that experience  $B$  will occur. The probability of  $B$ , given variable  $A$ , is given by the Law of Total Probabilities as  $P(B|A) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \dots + P(A_n)P(B|A_n)$ , where  $P(B|A_i)$  represents the conditional probability of  $B$ , given the state  $A_i$ . The conditional probability function relating the experience to the variable, that is,  $P(B|A)$ , is called the likelihood function (see “Model Description”). Using Bayes’s theorem, we can now estimate the posterior probability of any given state  $A_i$ , given the occurrence of the experience  $B$ , as

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(B)}.$$

In our analysis, we use the beta distribution (see “Model Description”) to describe both prior distributions and likelihood functions. The general form of the probability density function for a beta distribution is

$$P(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1 - x)^{\beta-1},$$

with  $0 \leq x \leq 1$  and  $P(x) = 0$  for all  $x$  outside this range, where  $\Gamma$  represents the Gamma function and  $\alpha$  and  $\beta$  are parameters of the beta distribution. Because this analysis involves two different beta distributions, we describe the prior distribution using a beta distribution in which  $\alpha = a$  and  $\beta = b$  and the likelihood function using a second beta distribution, in which  $\alpha = c$  and  $\beta = d$ .

For  $n = 100$ , for the prior distribution we can numerically approximate the probability  $P(A_i)$  by

$$P(A_i) = 0.01 p_A \left[ \frac{0.01(i - 1) + 0.01i}{2} \right] = 0.01 p_A(0.01i - 0.005),$$

where  $p_A(x)$  is the probability density function for the first beta distribution. Similarly, the probability of occurrence of the experience  $B$ , given the  $i$ th state  $A_i$ , can be computed as

$$P(B|A_i) = p_B \left[ \frac{0.01(i - 1) + 0.01i}{2} \right] = p_B(0.01i - 0.005),$$

where  $p_B(x)$  is the probability density function for the second beta distribution.

The posterior probabilities of each of the states can then be computed as follows. First, we compute the prior distribution. If we indicate the midpoint of the  $i$ th state as  $y_i = (0.01i - 0.005)$  and  $A_1, A_2, \dots, A_n$  are the  $n$  states



associated with variable  $A$ , then the prior probability distribution may be computed using the standard beta probability density function as

$$P(A_i) = 0.01p_A(y_i) = 0.01 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} (0.01i - 0.005)^{a-1} (1 - 0.01i + 0.005)^{b-1}.$$

Similarly, the conditional probability,  $P(B|A_i)$ , for each state can be computed using the beta probability function with parameters  $c$  and  $d$  as

$$P(B|A_i) = 0.01p_B(y_i) = 0.01 \frac{\Gamma(c+d)}{\Gamma(c)\Gamma(d)} (0.01i - 0.005)^{c-1} (1 - 0.01i + 0.005)^{d-1}.$$

Based on Bayes's theorem, the posterior probability of each state is given by  $P(A_i|B) = P(A_i)P(B|A_i)/P(B)$ , for  $i = 1, 2, \dots, n$ . Finally, we need to normalize these posterior probabilities, so that the sum  $P(A_1) + P(A_2) + \dots + P(A_n) = 1$ . Normalization is required because we have specified that  $A_1, A_2, \dots, A_n$  cover all possible states of the variable  $A$ , so that the sum of the probabilities of these states must add up to 1 for any prior or posterior distribution. We normalize the posterior probabilities by dividing each  $P(A_i|B)$  by the term

$$\sum_{i=1}^{100} P(A_i|B),$$

so that the final posterior probabilities are given by

$$\frac{P(A_i|B)}{\sum_{i=1}^{100} P(A_i|B)} = \frac{P(A_i)P(B|A_i)}{P(B)} \frac{P(B)}{\sum_{i=1}^{100} P(A_i)P(B|A_i)} = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^{100} P(A_i)P(B|A_i)}.$$

The posterior distribution generated by this procedure is not necessarily a beta distribution, even though the prior distribution was a beta distribution. However, this posterior distribution can be used as the prior distribution for the next experience because this posterior distribution is available as a computed distribution at the end of the process described above.

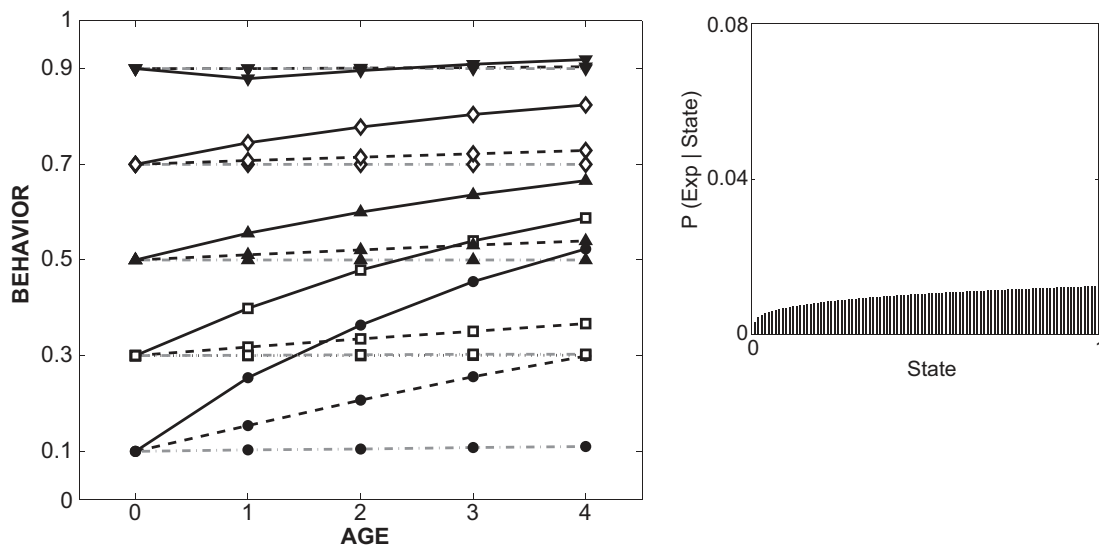
Although  $P(B)$  can take on many different values, the value of  $P(B)$  itself does not affect the computation of the posterior distribution, because of the process of normalizing the posterior probabilities (see equation above). An alternative method for computing the distributions for  $P(A_i)$  and  $P(B|A_i)$  would be to use the cumulative probability distribution function for a beta distribution. For example, the prior distribution can be computed as  $P(A_i) = \text{betaCDF}(0.01i, \alpha, \beta) - \text{betaCDF}(0.01(i-1), \alpha, \beta)$ , where  $\text{betaCDF}$  is the beta cumulative distribution function with parameters  $\alpha$  and  $\beta$ . This function is available in most standard statistical packages (e.g., R, SAS).

# Appendix C from J. A. Stamps and V. V. Krishnan, “Combining Information from Ancestors and Personal Experiences to Predict Individual Differences in Developmental Trajectories” (Am. Nat., vol. 184, no. 5, p. 000)

## Effects of Different Likelihood Functions on Developmental Trajectories Effects of the Reliability of Likelihood Functions on Developmental Trajectories

Differences among agents (individuals or genotypes) in developmental trajectories are most apparent for likelihood functions with intermediate reliabilities. When the likelihood function for a given experience is very unreliable (fig. C1), the experience has little effect on the behavior of any agent. As a result, differences across agents in developmental trajectories will be difficult to detect, and any behavioral differences observed in naive agents will be largely maintained through ontogeny. Conversely, when the likelihood function for a given experience is very reliable (fig. C4), the experience has a very strong effect on the behavior of all but those agents whose prior distributions have very low variance. As a result, after a relatively short period, most agents express similar behavior.

For all figures in this appendix, the means of the prior distributions (at age 0) are indicated by symbols—circles for 0.1, squares for 0.3, upward triangles for 0.5, diamonds for 0.7, and downward triangles for 0.9—and the variances of the prior distributions are indicated by lines—dot-dashed gray for 0.001, dashed black for 0.02, and solid black for the maximum variance for the given mean. The likelihood functions used to generate each set of developmental trajectories are indicated in the right-hand panel. For each likelihood function, we indicate the probability of the experience, given the state ( $P(\text{Exp}|\text{State})$ ) for each of the 100 states from 0 to 1.



**Figure C1:** Developmental trajectories (*left*) for a likelihood function with mean = 0.556, var = 0.076,  $\alpha = 1.25$ , and  $\beta = 1$  (*right*).

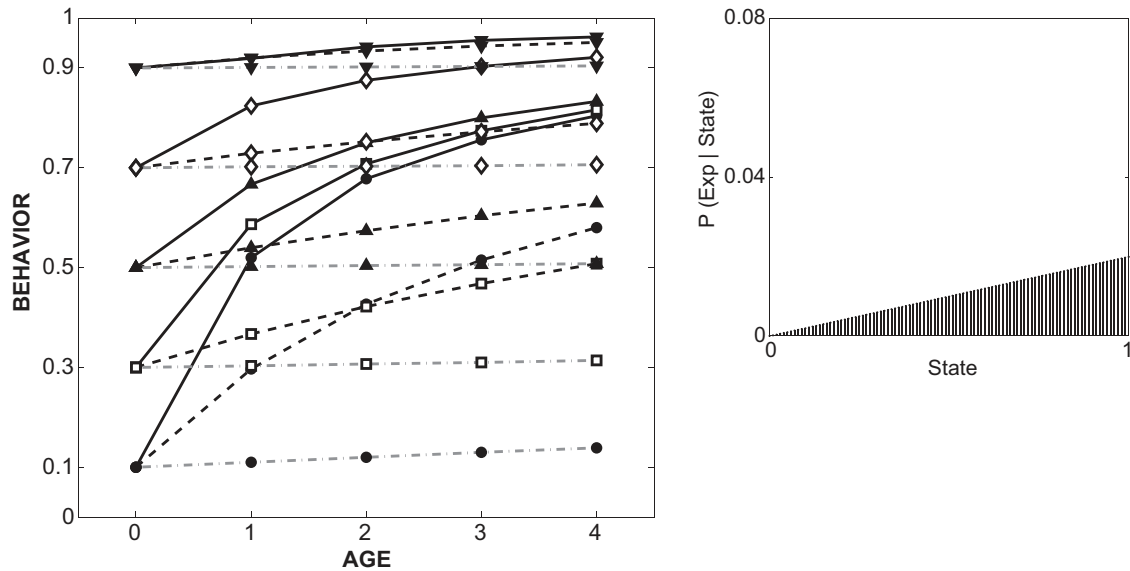


Figure C2: Developmental trajectories (left) for a likelihood function with mean = 0.667, var = 0.056,  $\alpha = 2$ , and  $\beta = 1$  (right).

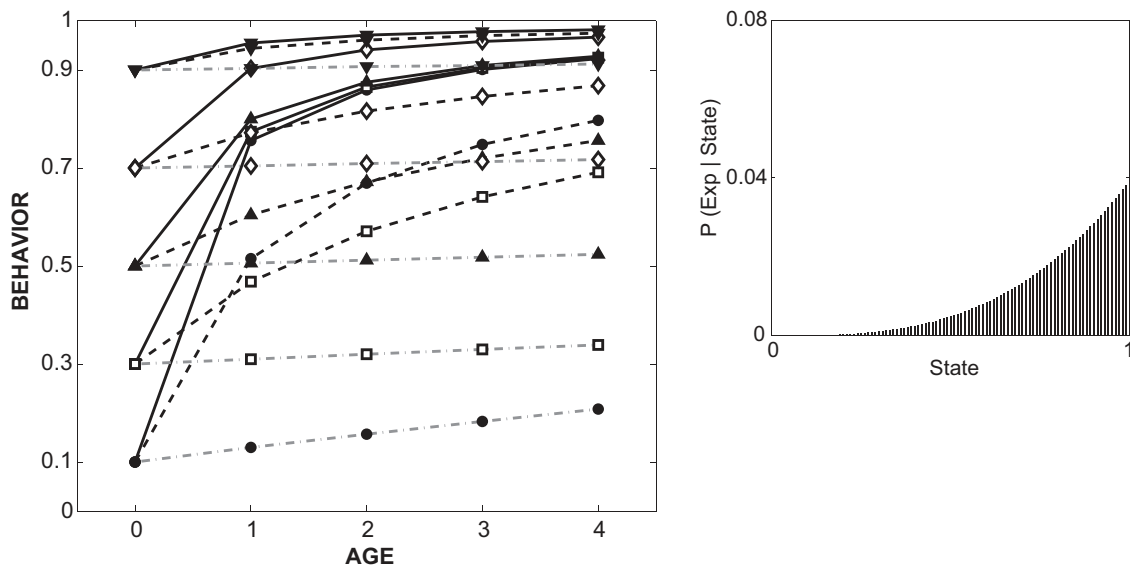
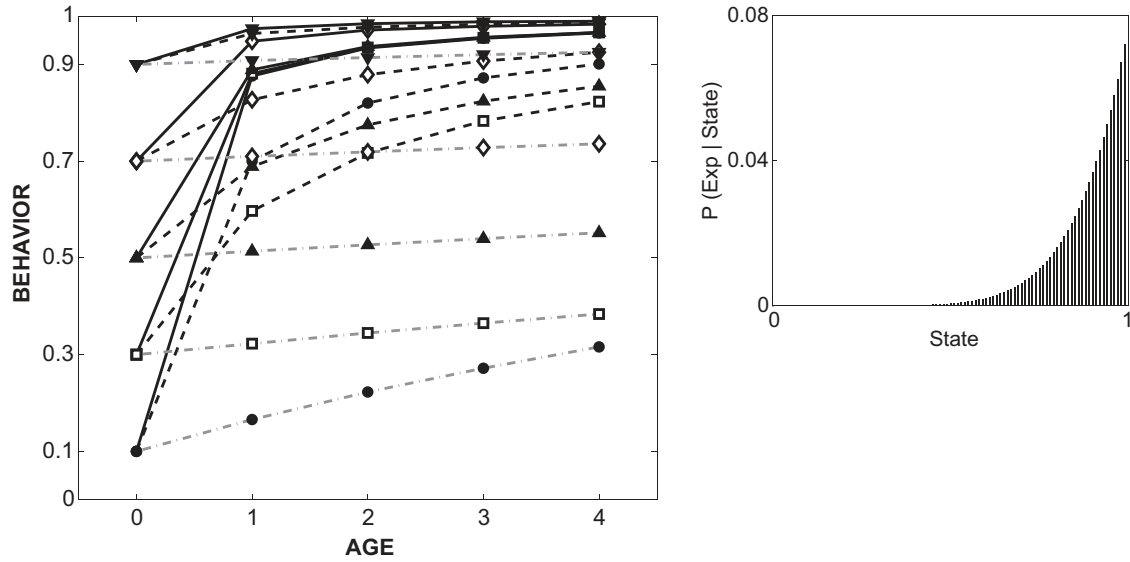


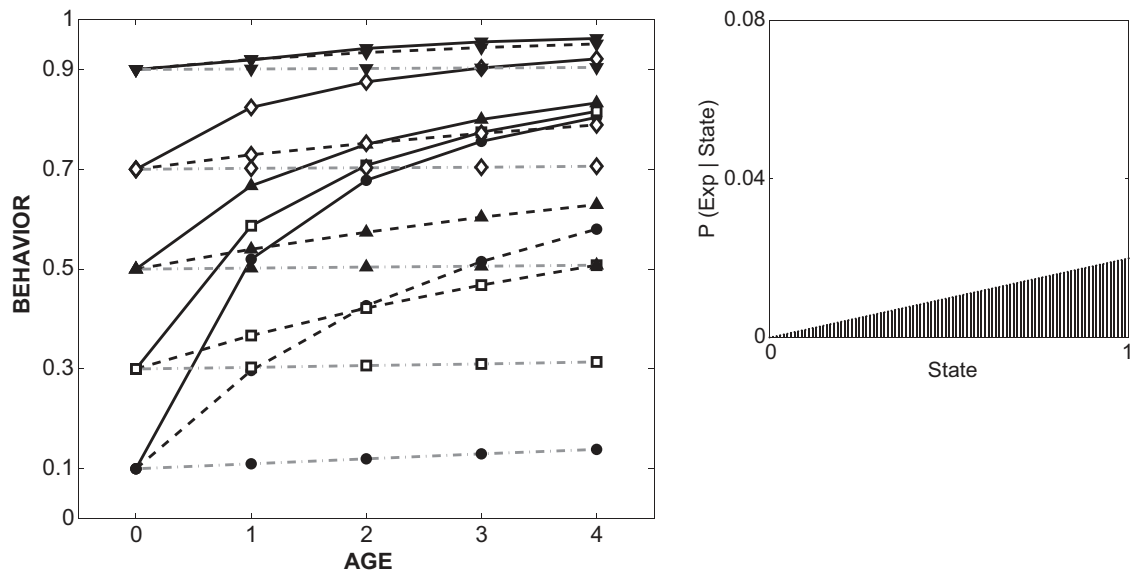
Figure C3: Developmental trajectories (left) for a likelihood function with mean = 0.800, var = 0.0267,  $\alpha = 4$ , and  $\beta = 1$  (right).



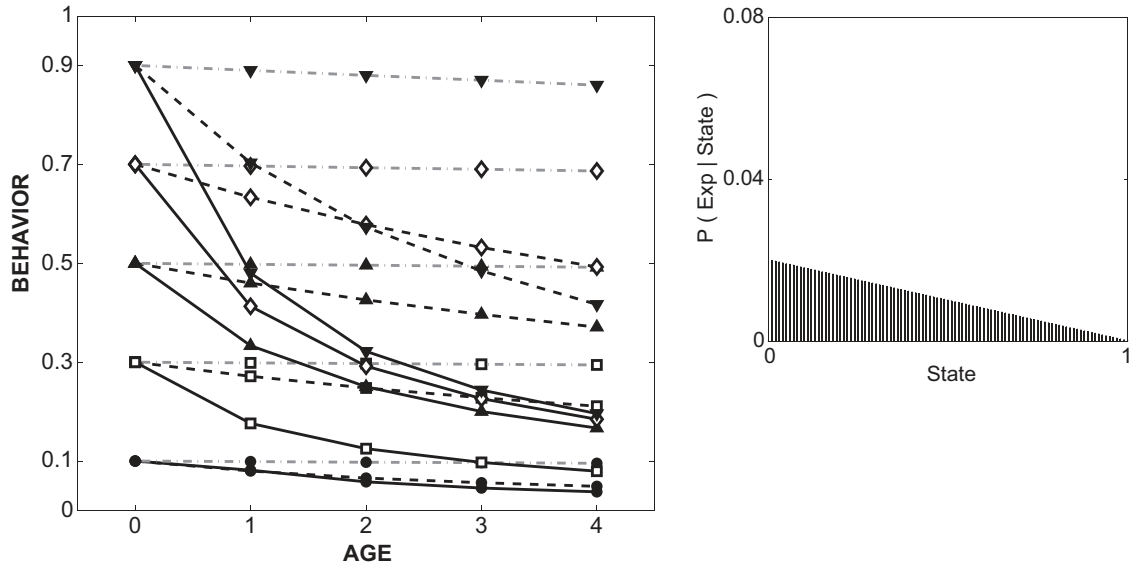
**Figure C4:** Developmental trajectories (*left*) for a likelihood function with mean = 0.889, var = 0.010,  $\alpha = 8$ , and  $\beta = 1$  (*right*).

### Effects of the Shapes of Likelihood Functions on Developmental Trajectories

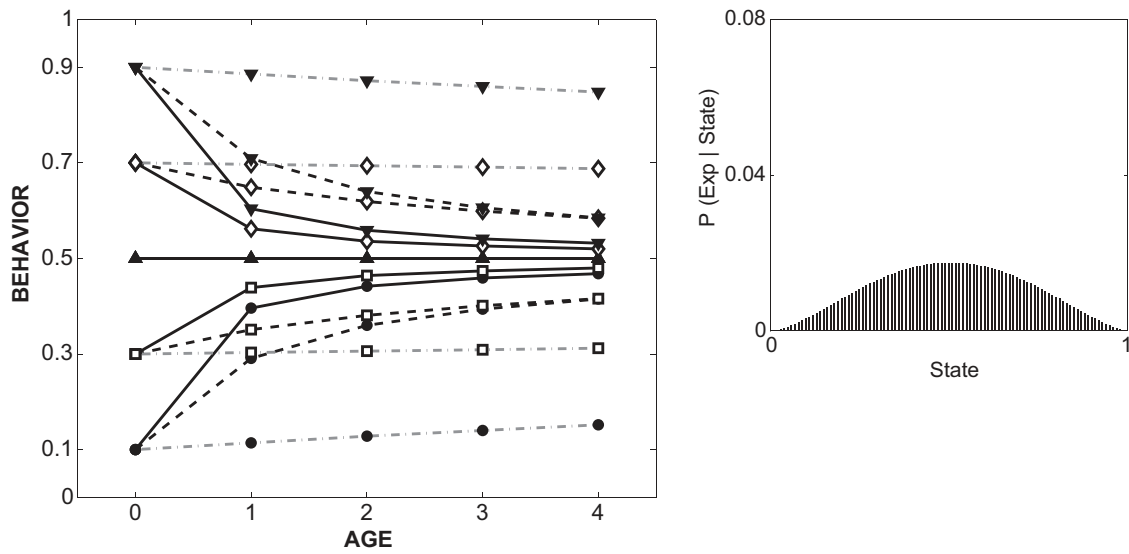
For likelihood functions of different shapes, the behavior of different agents tends to converge on the behavior encouraged by the experience. In addition, rates of change in behavior are higher early in ontogeny than later in ontogeny, differential consistency is lower early in ontogeny than later in ontogeny, differences among agents in prior distributions have long-lasting effects on behavior, and there are predictable relationships, across agents, between intercepts and the shapes or slopes of their developmental trajectories. These effects are illustrated here using a right-biased (fig. C5), a left-biased (fig. C6), and a unimodal (fig. C7) likelihood function.



**Figure C5:** Developmental trajectories (*left*) for a right-biased likelihood function with mean = 0.667, var = 0.056,  $\alpha = 2$ , and  $\beta = 1$  (*right*).



**Figure C6:** Developmental trajectories (*left*) for a left-biased likelihood function with mean = 0.333, var = 0.056,  $\alpha = 1$ , and  $\beta = 2$  (*right*).



**Figure C7:** Developmental trajectories (*left*) for a unimodal likelihood function with mean = 0.5, var = 0.04,  $\alpha = 2.625$ , and  $\beta = 2.625$  (*right*).

# Appendix D from J. A. Stamps and V. V. Krishnan, “Combining Information from Ancestors and Personal Experiences to Predict Individual Differences in Developmental Trajectories”

(Am. Nat., vol. 184, no. 5, p. 000)

## Effects of Different Response Functions on Developmental Trajectories

We model different response functions as follows: (1) a linear function,  $B = \mu$ ; (2) an asymptotic function,  $B = 1 - e^{-\mu^{0.3}}$ ; (3) an exponential function,  $B = (1/27)(e^{-\mu^{0.3}} - 1)$ ; and (4) a sigmoid function,  $B = 1/(1 + e^{10(\mu-0.5)})$ ;  $\mu$  is the mean of a prior or a posterior distribution and  $B$  is level of behavior. For each response function, we illustrate the developmental trajectories of 15 hypothetical agents, each of which was exposed four times to a personal experience with a right-biased likelihood function with intermediate reliability (mean = 0.67, var = 0.056,  $\alpha = 2$ , and  $\beta = 1$ ; see fig. C2).

For all figures in this appendix, the means of the prior distributions (at age 0) are indicated by symbols—circles for 0.1, squares for 0.3, upward triangles for 0.5, diamonds for 0.7, and downward triangles for 0.9—and the variances of the prior distributions are indicated by lines—dot-dashed gray for 0.001, dashed black for 0.02, and solid black for the maximum variance for the given mean. The shape of the response function is shown in a panel to the right of each set of developmental trajectories.

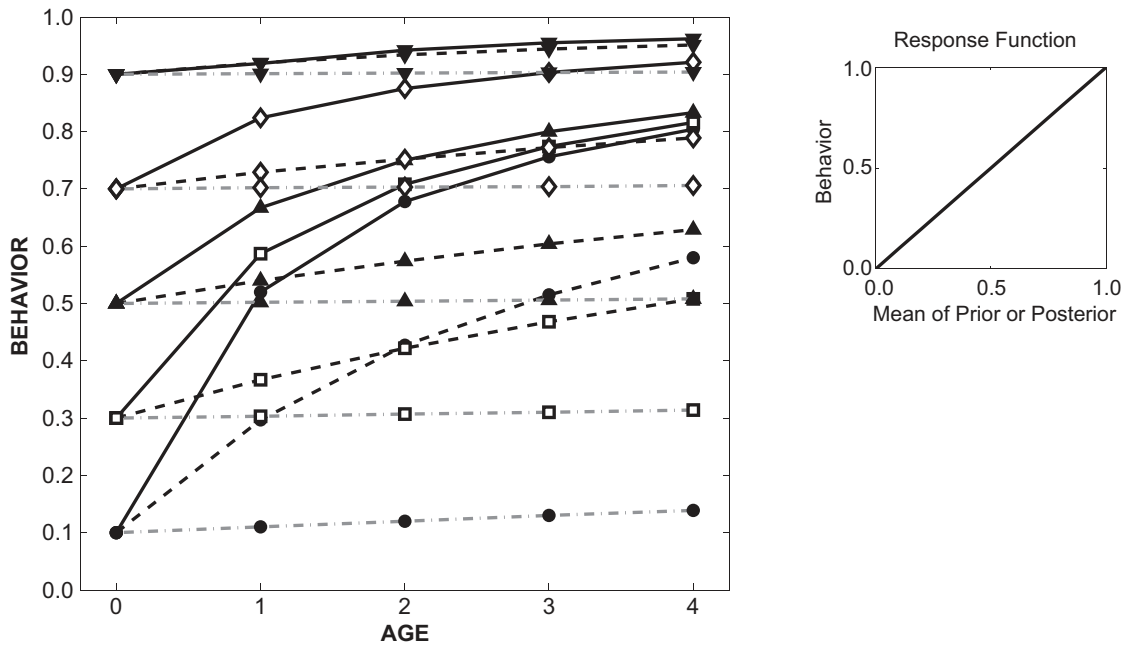


Figure D1: Developmental trajectories (*left*) for a linear response function (*right*).

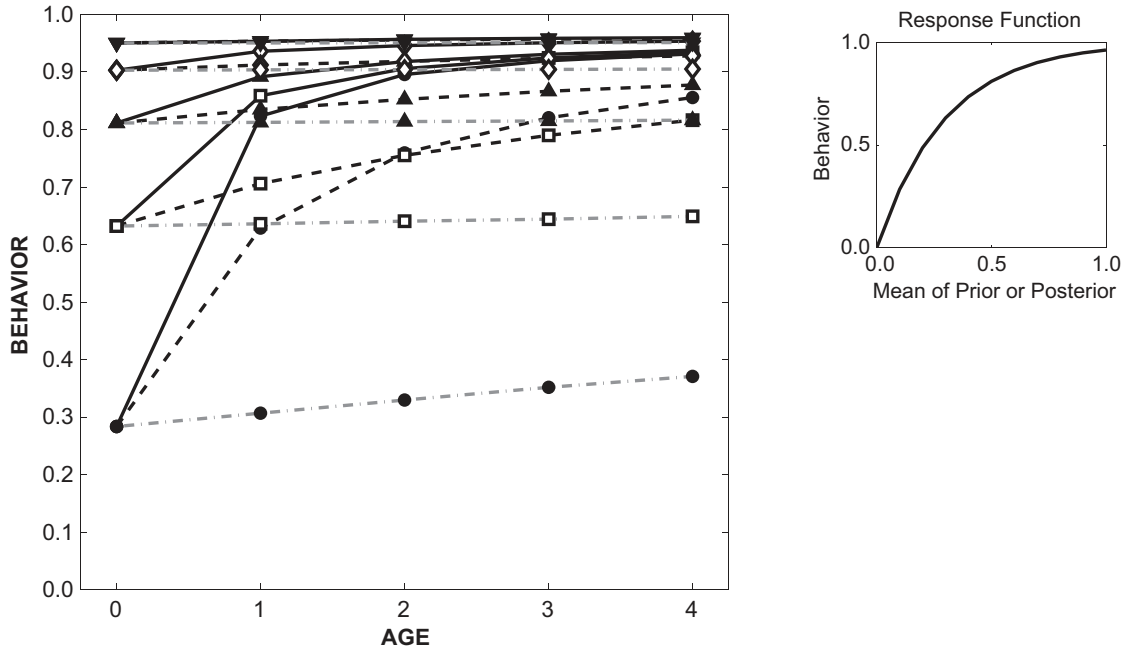


Figure D2: Developmental trajectories (*left*) for an asymptotic response function (*right*).

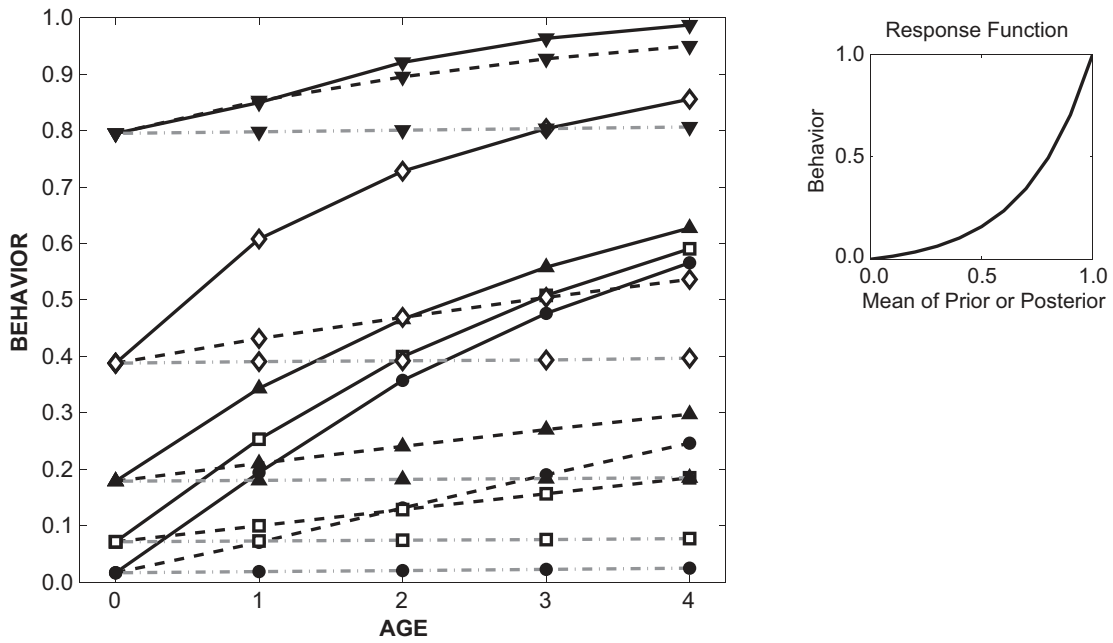


Figure D3: Developmental trajectories (*left*) for an exponential response function (*right*).

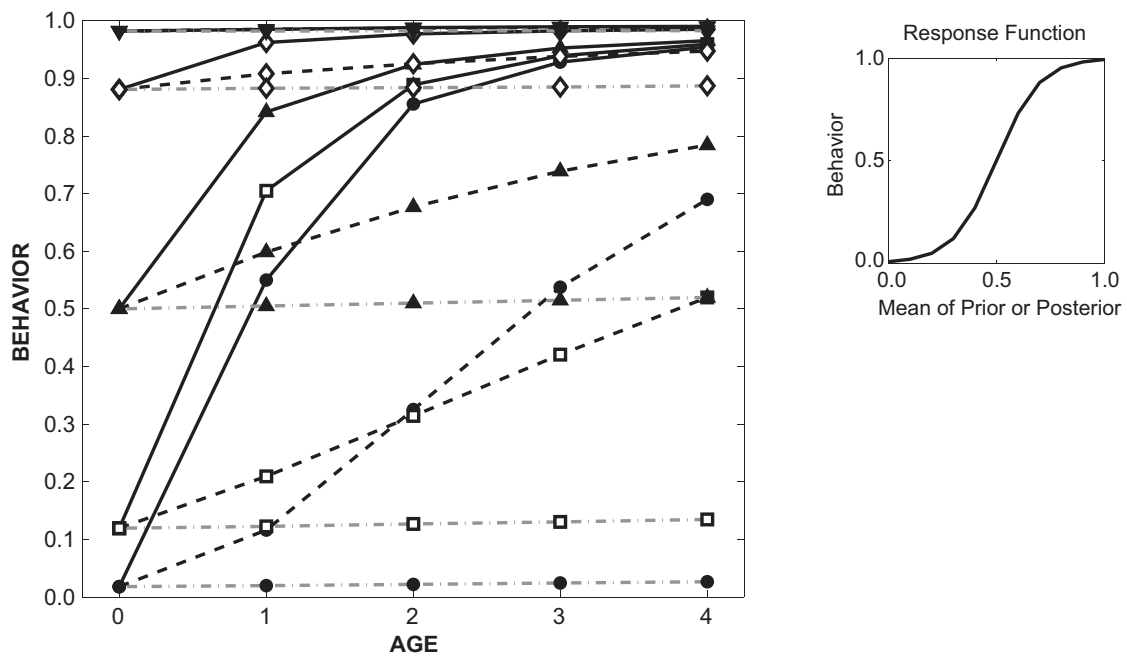


Figure D4: Developmental trajectories (*left*) for a sigmoid response function (*right*).



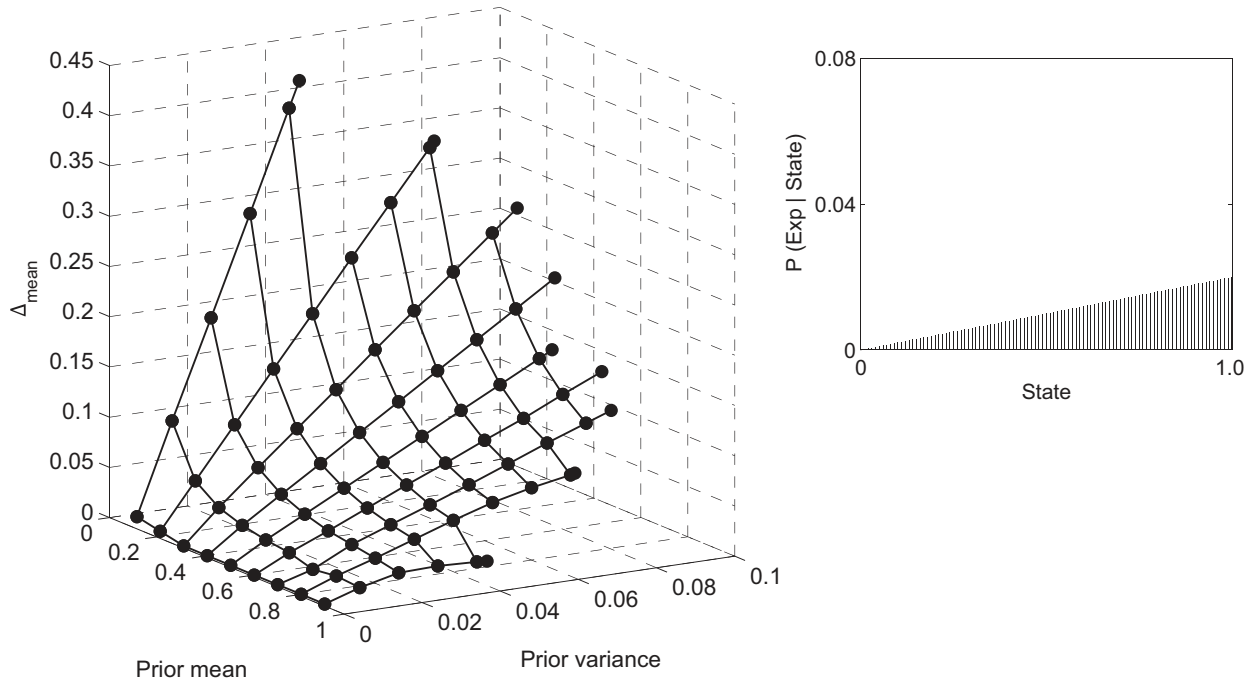
## **Appendix E from J. A. Stamps and V. V. Krishnan, “Combining Information from Ancestors and Personal Experiences to Predict Individual Differences in Developmental Trajectories” (Am. Nat., vol. 184, no. 5, p. 000)**

### **Joint Effects of Mean and Variances of Prior Distributions on Bayesian Updating**

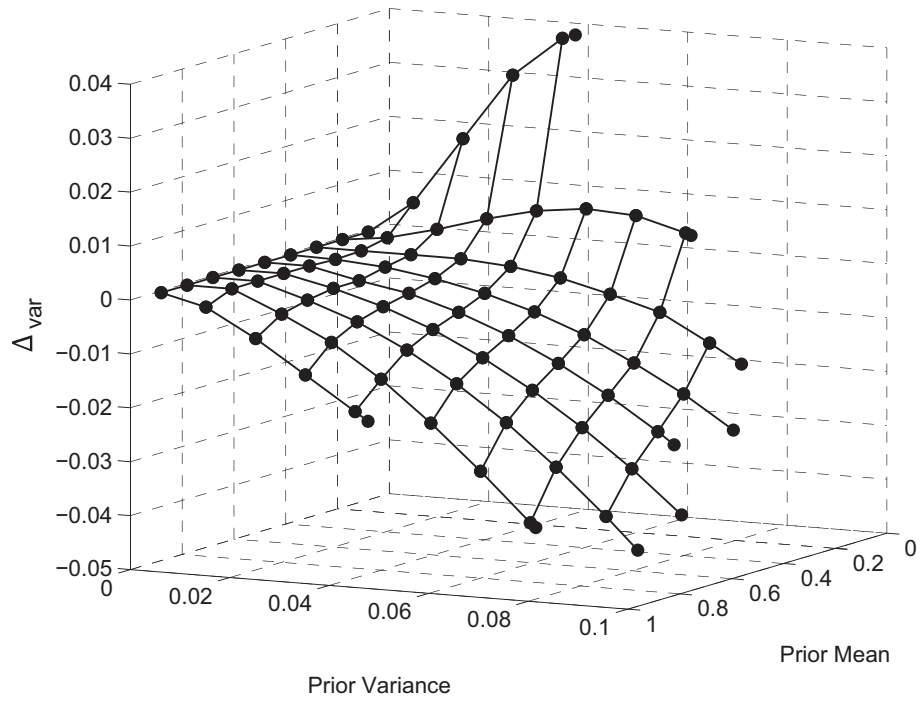
Here we consider how the means and variances of prior distributions jointly affect Bayesian updating, based on the model outlined in the text (also see app. B). We define a variable  $\Delta_{\text{mean}}$  to indicate the effect of a given experience on an individual’s estimate of the state, where  $\Delta_{\text{mean}} = (\text{mean of the posterior distribution} - \text{mean of the prior distribution})$ . A second variable,  $\Delta_{\text{var}}$ , indicates the effect of a given experience on an individual’s certainty about that estimate, where  $\Delta_{\text{var}} = (\text{variance of the posterior distribution} - \text{variance of the prior distribution})$ . Figures E1 and E2 indicate the values of  $\Delta_{\text{mean}}$  and  $\Delta_{\text{var}}$ , respectively, after individuals with a range of prior distributions have been exposed once to an experience with a right-biased likelihood function with intermediate reliability.

Many basic features of Bayesian updating are evident in figure E1. Because the likelihood function for this particular experience is right biased, the experience has a much stronger effect on the estimate of the state ( $\Delta_{\text{mean}}$ ) for individuals whose prior distributions had low mean values (e.g., prior mean = 0.1) than for individuals whose prior distributions had high mean values (e.g., prior mean = 0.9). In addition, regardless of their prior means, the same experience has less effect on the estimate of the state for individuals whose prior distributions had low variance (e.g., prior variance < 0.01) than for individuals whose prior distributions had higher variance.

The mean and variance of an individual’s prior distribution also determine how a given experience will affect an individual’s degree of certainty about its estimate of the state of the world, as indicated by  $\Delta_{\text{var}}$  (fig. E2). Although exposure to experience with a right-biased likelihood function usually reduces uncertainty (indicated by negative values of  $\Delta_{\text{var}}$ ), there are two important exceptions. First, if prior distributions have low mean values (here,  $\leq 0.2$ ), this experience increases rather than decreases uncertainty (indicated by positive values of  $\Delta_{\text{var}}$ ). In addition, this experience has little or no effect on uncertainty for individuals whose prior distributions had a low variance to begin with. Both of these patterns make intuitive sense. If the information provided by personal experience sharply conflicts with an individual’s belief before that experience, then the individual should be more uncertain about the state of the world after the experience than before. Otherwise, reasonably informative experience should reduce uncertainty, except for individuals whose uncertainty was already low before the experience.



**Figure E1:** Effects of an experience with a moderately reliable right-biased likelihood function on  $\Delta_{\text{mean}}$  (the difference between the mean of the posterior distribution and the mean of the prior distribution), for prior distributions with a range of means and variances. The likelihood function (*right*) indicates the probability of the experience, given the state, for each of the 100 possible states between 0 and 1. In this case, the likelihood function has a mean of 0.67 and a variance of 0.056; it was generated by a beta function in which  $\alpha = 2$  and  $\beta = 1$ .



**Figure E2:** Effects of a moderately reliable right-biased likelihood function on  $\Delta_{\text{var}}$  (the difference between the variance of the posterior distribution and the variance of the prior distribution), for prior distributions with a range of means and variances. The likelihood function is the same as that in figure E1. The axes for this figure are oriented differently from those in figure E1 to improve legibility.