A reaction norm framework for the evolution of learning: how cumulative experience shapes phenotypic plasticity

Jonathan Wright[1],[*] , Thomas R. Haaland[1],[2], Niels J. Dingemanse[3] and David F. Westneat[4]

[1]*Center for Biodiversity Dynamics (CBD), Department of Biology, Norwegian University of Science and Technology (NTNU), N-7491, Trondheim, Norway*
[2]*Department of Evolutionary Biology and Environmental Studies, University of Zürich, Winterthurerstrasse 190, CH-8057, Zürich, Switzerland*
[3]*Behavioural Ecology, Department of Biology, Ludwig-Maximilians University of Munich (LMU), 82152, Planegg-Martinsried, Germany*
[4]*Department of Biology, University of Kentucky, 101 Morgan Building, Lexington, KY, 40506-0225, USA*

ABSTRACT

Learning is a familiar process to most people, but it currently lacks a fully developed theoretical position within evolutionary biology. Learning (memory and forgetting) involves adjustments in behaviour in response to cumulative sequences of prior experiences or exposures to environmental cues. We therefore suggest that all forms of learning (and some similar biological phenomena in development, aging, acquired immunity and acclimation) can usefully be viewed as special cases of phenotypic plasticity, and formally modelled by expanding the concept of reaction norms to include additional environmental dimensions quantifying sequences of cumulative experience (learning) and the time delays between events (forgetting). Memory therefore represents just one of a number of different internal neurological, physiological, hormonal and anatomical 'states' that mediate the carry-over effects of cumulative environmental experiences on phenotypes across different time periods. The mathematical and graphical conceptualisation of learning as plasticity within a reaction norm framework can easily accommodate a range of different ecological scenarios, closely linking statistical estimates with biological processes. Learning and non-learning plasticity interact whenever cumulative prior experience causes a modification in the reaction norm (*a*) elevation [mean phenotype], (*b*) slope [responsiveness], (*c*) environmental estimate error [informational memory] and/or (*d*) phenotypic precision [skill acquisition]. Innovation and learning new contingencies in novel (laboratory) environments can also be accommodated within this approach. A common reaction norm approach should thus encourage productive cross-fertilisation of ideas between traditional studies of learning and phenotypic plasticity. As an example, we model the evolution of plasticity with and without learning under different levels of environmental estimation error to show how learning works as a specific adaptation promoting phenotypic plasticity in temporally autocorrelated environments. Our reaction norm framework for learning and analogous biological processes provides a conceptual and mathematical structure aimed at usefully stimulating future theoretical and empirical investigations into the evolution of plasticity across a wider range of ecological contexts, while providing new interdisciplinary connections regarding learning mechanisms.

Key words: behavioural plasticity, behavioural flexibility, learning rules, developmental plasticity, phenotypic equation, state-dependence, habituation curves.

CONTENTS

* Author for correspondence (Tel.: +47 91897210; Fax: +47 73596100; E-mail: jonathan.wright@ntnu.no).

I. INTRODUCTION

Natural systems contain many types of learning and memory, all of which appear to have evolved as adaptive responses to their ecological context and the specific challenges posed by particular forms of environmental variation (Stephens, 1991). For example, the learning involved in kin recognition imprinting (Bolhuis, 1991; Holmes & Mateo, 2007) or conditioned taste aversion to potentially toxic foods (Gustavson, 1977; Nicolaus & Nellis, 1987) is necessarily rapid and has potentially long-lasting adaptive effects irrespective of subsequent information or lack of it. By contrast, during classic habituation (see Shettleworth, 2010) or associative learning [e.g. when foraging (Hirvonen *et al.*, 1999; Stephens, Brown & Ydenberg, 2007)], there is a more gradual acquisition of information about often temporary conditions that can then be easily forgotten unless reinforced. A range of intermediate timeframes for learning and forgetting are also apparent, such as in the well-studied evolutionary ecology of spatial memory, food storing, territoriality and migration (see Healy & Hurly, 2004; Shettleworth, 2010), and song learning in birds (Catchpole & Slater, 1995). Therefore, the types of salient environmental cues that animals attend to, how quickly experience of those cues affects behaviour and how long such effects persist, appear to make sense according to the particular environmental factor of importance, its predictability and its rate of change relative to the lifetime of the individual. Such obviously adaptive patterns of phenotypic change due to learning (and forgetting) in natural populations argue for the development of an evolutionary framework promoting the scientific understanding of learning in an ecological context.

The biological literature contains an encouraging number of recent studies on the evolution of learning [see Mery & Burns, 2010; Fawcett, Hamblin & Giraldeau, 2012 and accompanying commentaries; Greggor *et al.*, 2019]. Theoretical models of behavioural learning [see Feldman & Aoki (2014) and other papers in this special issue] have explored factors affecting the evolution of adaptive learning rules in specific contexts ranging from foraging (McNamara, 1985; McNamara & Houston, 1987; Stephens *et al.*, 2007; Eliassen *et al.*, 2009) to mating strategies (Dukas, Clark & Abbott, 2006) and other game-theoretical issues involving frequency-dependent social interactions (Hamblin & Giraldeau, 2009; Dubois, Morand-Ferron & Giraldeau, 2010; Katsnelson *et al.*, 2012; Afshar & Giraldeau, 2014; Lee *et al.*, 2016; Aplin & Morand-Ferron, 2017; McNamara & Leimar, 2020). However, despite some excellent conceptual reviews of learning and other cognitive processes within the field of animal behaviour (e.g. Dukas, 2004, 2013), a more complete theoretical framework that links a common set of model parameters with the biological processes involved in adaptive learning has yet to be fully developed.

Similarly, there are excellent empirical studies on the evolution of learning in specific contexts, including elegant selection experiments on the role of learning in oviposition decisions in *Drosophila* by Mery & Kawecki (2002) and Dunlap & Stephens (2009), which highlight the importance of learning in response to temporally auto-correlated environments. Some studies have also attempted to quantify the fitness costs and benefits in specific cases of learning (e.g. Johnston, 1982; Sullivan, 1988; Hollis *et al.*, 1997; Mery & Kawecki, 2004; Mahometa & Domjan, 2005), and provided suggestive evidence for adaptive differences in learning between populations, sexes or species (e.g. Simons *et al.*, 1992; Balda, Kamil & Bednekoff, 1996; Lefebvre, 1996; Jackson & Carter, 2001; Dunlap *et al.*, 2006; see also Shettleworth, 1984, 2010). The taxonomically widespread diversity of forms of learning and the somewhat disconnected nature of most of the theoretical and empirical literature combine to suggest a need for a conceptual framework for learning that is fully embedded within evolutionary biology.

Any general framework for the evolution of learning (and forgetting) would benefit from being incorporated within the study of phenotypic plasticity. Learning clearly meets the broad definition of plasticity, which is variation in the phenotype expressed by a given genotype or individual due to variation in the environment (Schlichting & Pigliucci, 1998; Scheiner, 2006). Learning can be defined as a meaningful change in behaviour with individual experience (Shettleworth, 1984, 2010; Stephens, 1989; Dukas, 1998;

Dall *et al.*, 2005; Mery & Burns, 2010; Staddon, 2016). The relevant experience in almost all cases of learning is the result of some prior exposure to the environment, broadly defined, and so is a form of phenotypic plasticity (see also Dukas, 2004, 2013). Considering learning explicitly as a form of plasticity has several useful consequences. First and foremost, there is considerable evolutionary theory concerning the evolution of plasticity (see Schlichting & Pigliucci, 1998; Scheiner, 2006; Botero *et al.*, 2015; Tufto, 2015), which can usefully be applied to learning. Integrating learning into the current framework for phenotypic plasticity also helps clarify similarities and differences between learning as a form of plasticity *versus* plasticity that does not involve learning. Finally, prior experience may interact with non-learned responses to environmental factors. That is, phenotypically plastic responses to environmental variation might differ depending upon the degree of learnt experience, but there are few suggestions in the literature about how and when this might occur. Here we outline a conceptual framework for integrating learning into evolutionary theory on phenotypic plasticity that is specifically designed to tackle such issues.

In evolutionary biology, phenotypic plasticity is often represented in terms of norms of reaction (see Kawecki & Stearns, 1993; Scheiner, 1993; Schlichting & Pigliucci, 1998; Nussey *et al.*, 2005; Dingemanse *et al.*, 2010), although sometimes a character-state approach may be preferable (Via *et al.*, 1995). Reaction norms are functions relating the phenotypic values for a single genotype or individual to a given set of environmental conditions (internal, external, abiotic, biotic or social). These functions allow us to think explicitly about the form of plasticity, its variation within and among populations, the genetic and environmental sources of that variation, and the selective forces acting upon different components of plasticity in terms of the parameters defining the elevation and slope (or more complex non-linear shapes) of a reaction norm (DeWitt, Sih & Wilson, 1998; Ghalambor *et al.*, 2007; Nussey, Wilson & Brommer, 2007; Lande, 2009; Murren *et al.*, 2015). Reaction norms can encompass any functional relationship between an environmental axis and the phenotype. Doing the same for learning therefore allows us to utilise all of the analytical and theoretical tools that have been developed for investigating the evolution of non-learning plasticity (see Nussey *et al.*, 2007; Stinchcombe & Kirkpatrick, 2012; Scheiner, 2013; Chevin & Lande, 2015). However, learning is a form of plasticity with particular properties – it is a change in behaviour in response to a cumulative sequence of environmental cues. So, rather than a phenotype simply responding plastically to a particular axis of environmental variation, with learning the phenotype responds differently to each environmental stimulus depending upon whether it has been experienced before; thus for learning the cues used to adjust the phenotype consist of a temporally ordered series of exposures to the environment. Such reaction norm representations of learning based upon sequences of prior exposures or experiences are well illustrated by habituation curves from learning psychology (Bills, 1934; Jaber, 2011). The learning rules (or cognitive mechanisms) thought to govern the shapes of such learning curves can thus be equated with behavioural 'rules-of-thumb' (McNamara & Houston, 1980) or 'strategy sets' (*sensu* Grafen, 1984) used to produce adaptive behavioural reaction norms. Learning reaction norms can therefore capture theoretical expectations from mathematical models for the evolution of optimum learning rules (e.g. Lotem & Halpern, 2012; Aoki & Feldman, 2014; McNamara & Leimar, 2020), because they are simply mathematical functions that define the effects of those learning rules on the phenotype.

Another useful consequence of integrating learning and phenotypic plasticity concerns the responses of non-behavioural traits to cumulative environmental exposures and prior experiences, often over much longer timescales. In this regard, learning (and the use of memory) bears a notable resemblance to cumulative environmental effects on early-life development and aging during the lifetime of the organism (West-Eberhard, 2003; van de Pol & Verhulst, 2006; Nussey *et al.*, 2007; Wolpert *et al.*, 2011; Stamps & Frankenhuis, 2016), acquired immunity in mammals (Janeway *et al.*, 2005; Flajnik & Kasahara, 2010), acquired responses of phytophagous insects to plant defences (Papaj & Prokopy, 1989; Bernays & Chapman, 1994), physiological acclimation (Angilletta, Niewiarowski & Navas, 2002; Schulte, Healy & Fangue, 2011; Seebacher, White & Franklin, 2015), and the cumulative hormonal modulation of behaviour (Hsu, Ryan & Wolf, 2006; Oliveira, 2009). Thus, in addition to learning and cognition affecting behaviour, there appears to be a range of physiological and developmental processes driving phenotypic plasticity based upon the effects of accumulated environmental exposures over a range of different biological timescales. Interestingly, for these responses to be adaptive, some degree of temporal autocorrelation in environmental conditions and/or fitness pay-offs is necessary, where prior experience usefully informs current expectations. Hence, these processes are characterised by gradual and cumulative (and often non-linear) changes to some important aspect or property of the organism, such as energy reserves, somatic growth or informational memory. Such properties of the individual can be conceptualised as internal 'state' variables that carry over their values and fitness consequences from one decision (or time) event to another, and the optimal strategy across a sequence of events can then be explored using state-dependent models (see Houston & McNamara, 1999; Clark & Mangel, 2000; Dall & Johnstone, 2002; Dunlap & Stephens, 2012). Thus, state-dependent behavioural plasticity not normally associated with learning (e.g. mass-dependent foraging in birds) can also involve *cumulative* environmental effects. Likewise, in some cases it might be difficult to distinguish learning (and memory) from other ongoing cumulative organismal responses to the environment, such as physiological adjustments or anatomical development. So, although here we focus upon examples of behavioural learning involving neurological mechanisms, we suggest that learning is not entirely unique and can be envisioned more generally as plasticity in response to an ordered series of environmental exposures. By combining various

theoretical and statistical approaches to learning, state-dependence and other analogous biological processes, it is possible to construct a single overarching conceptual framework for the evolution of any reaction norm involving cumulative environmental effects on phenotypic change over the lifetime of an organism, whether it is learning, developmental plasticity, acquisition and acclimation, as well as diurnal, seasonal and age effects.

Learning and other analogous biological phenomena (e.g. development, acclimation, etc.) influence a wide range of phenotypic traits and their functional importance is hard to overstate, including their effects on the rates and directions of phenotypic evolution (West-Eberhard, 2003, 2005; Brown, 2013; Dukas, 2017; Dayan et al., 2019). In this paper, we outline a general conceptual framework for such cumulative environmental effects on plasticity, with graphical and mathematical representations for learning that are fully embedded within wider evolutionary biology theory. This reaction norm approach has several useful advantages in that it allows us to study and understand learning using the evolutionary and statistical tools previously only applied in the context of non-learning phenotypic plasticity. Importantly, since prior cumulative exposure to an environment can constitute a reaction norm axis that involves 'learning', and this axis could also influence the nature of the reaction norm in other non-learning environment dimensions, we focus some attention on learning in the context of multidimensional plasticity (sensu Westneat et al., 2019) in which the phenotype responds plastically to two or more environmental variables. In doing so, we classify the different ways that learning can alter conventional non-learning reaction norms, explain the utility of formally defining the duration of 'time events' within any study, and demonstrate the importance of temporal autocorrelations in the environment for the evolution of learning via an individual-based simulation model. Our overarching aim is to facilitate the development of predictive hypotheses concerning the evolution of learning (and memory) and the empirical quantification of variation in learning (and other analogous biological phenomena) at key biological levels of interest (within and among individuals, genotypes and species). The reaction norm approach fully incorporates the statistical models that provide operational links to data analysis, and therefore is a means for conceptually matching statistical estimates to specific biological hypotheses. Here we focus largely upon the adaptive consequences of any phenotypic adjustments that might result from learning and other analogous processes (in plants as well as animals), rather than the specific types of cognitive, physiological or morphological mechanisms by which those phenotypic adjustments are achieved. This is not to say that cognitive mechanisms or evolutionary histories are unimportant (see Fawcett et al., 2012), but rather that unencumbered thinking about the ultimate function of phenotypic plasticity has made it possible to understand more fully the general principles behind its evolution, and the same is likely to be true at this initial stage for the evolution of learning and other cumulative plasticities. Contextualising learning as a specific

aspect of phenotypic plasticity using a reaction norm framework should offer new insights and avenues for research regarding the evolution of phenotypic plasticity and adaptive 'learning' (in the broadest possible sense), as well as the processes and mechanisms involved.

II. LEARNING AS A TYPE OF PLASTICITY: A REACTION NORM PERSPECTIVE

The concept of the reaction norm is well established within evolutionary biology (see Kawecki & Stearns, 1993; Scheiner, 1993; Schlichting & Pigliucci, 1998), being used in both theoretical and empirical investigations of adaptive phenotypic plasticity, including in the context of behavioural and life-history variation (see Nussey et al., 2007; Dingemanse et al., 2010; Westneat et al., 2011; Westneat, Wright & Dingemanse, 2015). Given the parallels (and sometimes confusion) between plasticity and learning, we suggest a formal extension in the form of reaction norms in which phenotypic expression occurs in response to a sequence of cumulative environmental experiences (see Section I). However, expanding the reaction norm approach to include learning (and forgetting) and other analogous biological phenomena means that we first need to consider key methodological issues that have already arisen as part of reaction norm studies of phenotypic plasticity.

Phenotypic plasticity has been categorised in many ways, and one dichotomy of interest here is between 'irreversible' (or 'developmental' or 'organisational') plasticity in traits that are usually expressed only once, such as the determinate growth of morphological characters or nervous systems, versus 'reversible' (or 'contextual' or 'activational') plasticity in traits that are expressed repeatedly with various different values throughout a lifetime, such as behaviour and hormones (Agrawal, 2001; West-Eberhard, 2003; Nussey et al., 2007; Dingemanse et al., 2010; Ord, Stamps & Losos, 2010; Snell-Rood, 2013; Nelson & Kriegsfeld, 2017). In reality, most instances of phenotypic plasticity fall between these two extremes in being more or less reversible in response to more or less long-lasting environmental effects (see Piersma & Drent, 2003), and this is perhaps particularly true of the various timeframes observed in learning and forgetting (see Section VII) and other analogous biological phenomena (see Section I). Nevertheless, this dichotomy has been instructive in highlighting the range of timescales and degrees of reversibility in the expression of phenotypes, and the diverse influence of environments underlying different forms of plasticity. Different plasticities tend to operate over different timeframes within the same lifetime, and these can interact in their effects on phenotypic expression. For example, developmental plasticity early in life may affect subsequent levels of reversible plasticity or behavioural responsiveness (Dingemanse et al., 2010; Stamps & Groothuis, 2010; Dingemanse & Wolf, 2013), and this itself has been a precept for many studies

over the years in behavioural endocrinology (Becker *et al.*, 2002; Nelson & Kriegsfeld, 2017).

As we will show, this interaction between the effects of different environmental axes, or multidimensional plasticity (*sensu* Westneat *et al.*, 2019), is particularly relevant for learning (and forgetting) and other analogous biological phenomena involving cumulative environmental effects, because learning often interacts with various aspects of the reaction norms describing plasticity without learning (see Section V). Therefore, the single environmental (*x*-axis) dimension used in many learning and non-learning plasticity reaction norms is a simplification, mostly for ease of understanding. The different timeframes of multiple environments combining to affect the same phenotypic trait, plus the obvious complexity of natural environments, means that most natural reaction norms are in effect multidimensional as organisms respond simultaneously to multiple (interacting) environmental factors (Westneat *et al.*, 2019).

Most reaction norms are also pragmatically presented as linear in the first instance, if only for simplicity, when we know that natural cases of phenotypic plasticity can involve non-linearities (Gavrilets & Scheiner, 1993; Murren *et al.*, 2014; Beaman, White & Seebacher, 2016; Arnold, Kruuk & Nicotra, 2019). Indeed, reaction norms involving cumulative experience are especially likely to be non-linear [e.g. exponential learning or habituation curves (Bills, 1934; Shettleworth, 2010; Jaber, 2011)]. This is because of Bayesian and other non-linear functions that can be used to describe the gradual cognitive and physiological processes that sequentially change the phenotype or (informational) state of the organism during the course of cumulative environmental exposures or prior experiences, usually in some curvilinear fashion until an upper maximum or lower minimum limit is reached (for parallel arguments concerning development, see Stamps & Frankenhuis, 2016). Therefore, greater insight will always be gained if the non-linear function fitted in the statistical model corresponds to a particular hypothesis or theoretical expectation regarding the particular biological mechanism or learning rule involved (see Lotem & Halpern, 2012; Aoki & Feldman, 2014; McNamara & Leimar, 2020; Westneat *et al.*, 2020). Oversimplifying the biology and using linear reaction norms to represent non-linear processes can lead to inappropriate conclusions from the application of incorrect statistical models in empirical studies. Luckily, the complexities of multidimensional and most forms of non-linear learning reaction norms can be accommodated into existing statistical models.

The statistical procedures involved in analysing reaction norms of repeatedly expressed traits also have the potential to modify the exact meaning of the reaction norm intercept and slope (Westneat *et al.*, 2020). For example, some reaction norms may use a zero value along the environmental axis to denote a true value of a ratio scale (e.g. a chemical concentration), but unbalanced data, interval scales or multidimensional reaction norms may make it preferable to mean-centre the zero of the reaction norm *x*-axis to represent the

average environment experienced (see van de Pol & Wright, 2009; Dingemanse *et al.*, 2010). Reaction norms involving an *x*-axis of cumulative experience instead would reasonably always place the zero value and thus intercept at the first instance of a stimulus, equating to when the individual has no prior experiences. Whether the reaction norm describes learning *per se* or an analogous cumulative phenomenon, this 'left-centring' captures the sequential and ordered nature of the organism's exposures to the environmental axis, which differs from that of other forms of plasticity in which the phenotype reacts to any given order and position of exposures along the environmental axis. We focus on applying the reaction norm approach explicitly to cases of learning, but we note that many of these ideas could apply to plasticity involving other sorts of cumulative experience. We occasionally transition to make this point explicitly.

III. A GRAPHICAL DESCRIPTION OF LEARNING REACTION NORMS

A conventional reaction norm can be depicted graphically (Fig. 1A), with the elevation representing the mean phenotypic value expressed by the focal individual in its average mean-centred environmental condition, and the (in this case) linear slope of the reaction norm representing the individual's responsiveness or degree of phenotypic change per unit of environmental change (Nussey *et al.*, 2007; Dingemanse *et al.*, 2010). This unidimensional depiction can be expanded to a multidimensional reaction norm surface (Fig. 1B) by including more than one environmental *x*-axis (Westneat *et al.*, 2019). Figure 1C shows how the curvilinear effects of learning can be similarly illustrated in an individual reaction norm plot for simple unidimensional cases, such as habituation. This is achieved by plotting the timing of the individual's cumulative exposures or prior experiences along a single environmental *x*-axis. Therefore, learning plasticity differs from non-learning plasticity in the use of a left-centred environmental axis depicting the level of prior individual experience (Fig. 1C), as opposed to independent exposures to environmental values with no particular order (Fig. 1A).

While unidimensional depictions of learning (Fig. 1C) are common in laboratory studies of behaviour, most learning occurs in complex natural environments and multidimensionality is likely a part of most learning reaction norms. Interactions between the effects of prior exposures and one or more environmental gradients (Fig. 1D) are likely common and are central to any understanding of the evolutionary forces shaping learning. For example, social foragers on ephemeral clumps of food tend to switch between searching for new clumps ('producing') and joining conspecifics at already discovered clumps ('scrounging') in a negatively frequency-dependent way that depends conditionally upon the current proportion of producers *versus* scroungers in the local population (i.e. the social environment represented on the *x*-axis of a non-learning reaction norm;
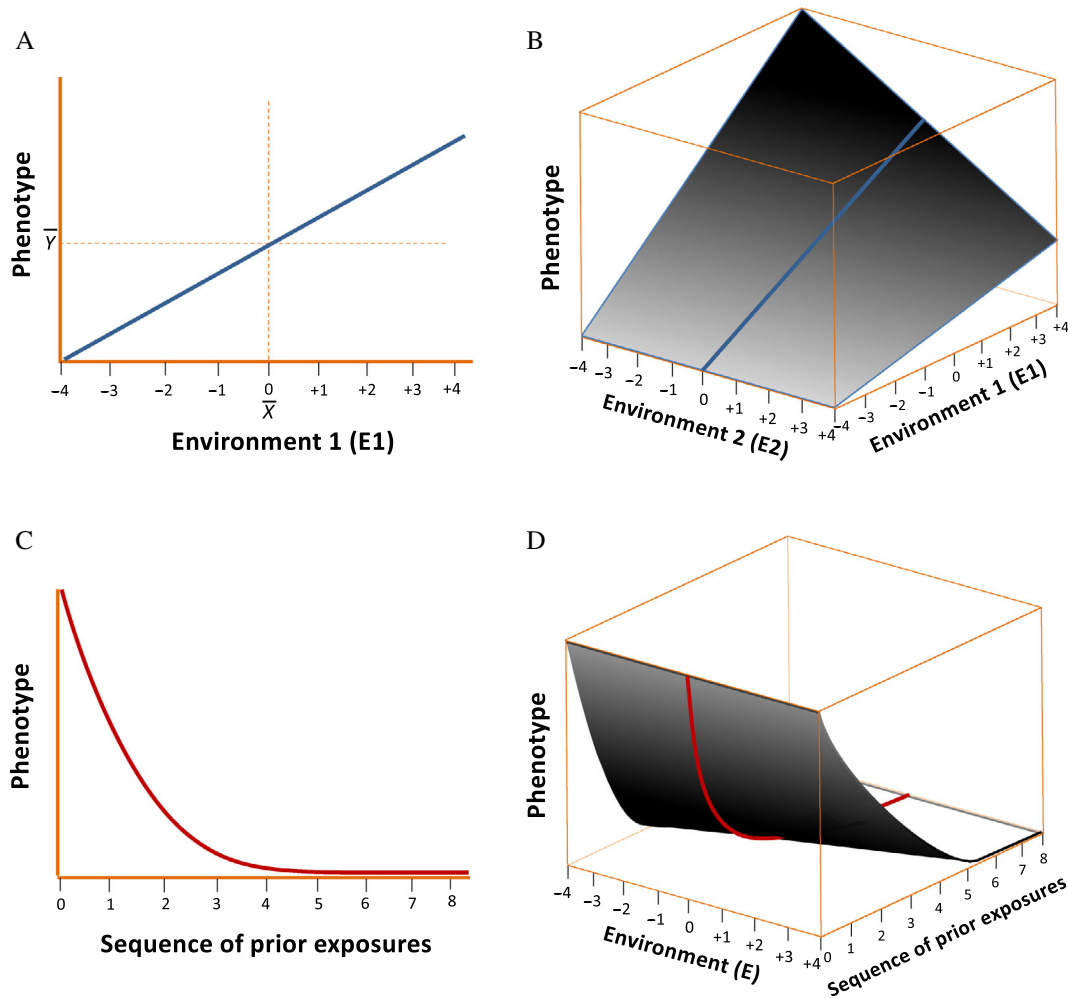
Fig. 1. Illustrations of a single individual's unidimensional and multidimensional reaction norms for non-learning phenotypic plasticity in response to environmental variation (blue reaction norms), and learning plasticity as a result of a cumulative sequence of prior exposures (red reaction norms). (A) Non-learning unidimensional plasticity as a linear response to the mean-centred environmental variable, E1 (e.g. foraging effort with increasing prey profitability), with the elevation $\overline{Y}$ representing the mean phenotypic value for that individual in its average environmental condition $\overline{X}$. (B) Non-learning multidimensional plasticity in response to two environmental variables, E1 and E2, with an interaction between them producing a warped reaction norm surface (e.g. predation threat and the need for vigilance moderating the positive effect of prey profitability on foraging effort). (C) Learning unidimensional plasticity following a particular sequence of evenly spaced prior exposures in which the behaviour decreases non-linearly – i.e. the effect per exposure declines with increasing prior experience (e.g. exponential effects of habituation to a benign novel object near a food source). (D) Learning multidimensional plasticity with the effect of cumulative experience from a sequence of events interacting in response to some additional environmental effect, E (e.g. habituation to a benign novel object taking longer with an increasing perceived predation threat). The blue and red lines thus represent unidimensional reaction norms in A and C, and reaction norm surface values at the mean-centred (zero) values of the environmental axes E2 in B and E in D. The darker shading of the grey reaction norm surfaces represents higher phenotypic values in B and later phenotypic expressions in D. See main text for more details, but note that the particular cases here were chosen for the purposes of illustration. In real systems, non-learning plasticity reaction norms A and B can also be non-linear, whilst learning reaction norms C and D can be linear, and both may involve more than two (x-axis) environmental effects.

Barnard & Sibly, 1981). House sparrows (*Passer domesticus*) show additional evidence of learning in such scenarios, based upon prior experience of cues relating to the pay-offs of producing *versus* scrounging early in life and/or during previous foraging sessions (Katsnelson *et al.*, 2008; Belmaker *et al.*, 2012). Exactly how these different forms of non-learning and learning plasticity interact and with what fitness consequences is likely to be a rich area for further study.

For the purposes of illustration, Fig. 1 shows the reaction norm for only a single individual. However, with sufficient data for multiple individuals one can estimate

among-individual variation in the elevations and slopes of both non-learning and learning reaction norms [e.g. in neophobia, habituation and novel cue learning in foraging house sparrows (Ensminger & Westneat, 2012; Moldoff & Westneat, 2017)]. With additional information on genetic relatedness, we could also use quantitative genetic approaches to partition this individual variation into the heritable *versus* permanent environmental (developmental) effects on elevations and plasticity (see Nussey *et al.*, 2007). This has been done for unidimensional learning reaction norms involving habituation in the form of decreases in exploration activity as a result of successive exposures to the same benign novel environment, thereby showing significant genetic variation in these learning reaction norm intercepts and slopes in both three-spined sticklebacks (*Gasterosteus aculeatus*; Dingemanse *et al.*, 2012*a*) and great tits (*Parus major*; Dingemanse *et al.*, 2012*b*). Log-transformations allow such exponential effects to be modelled as simple linear functions (e.g. Dingemanse *et al.*, 2012*a*; Moldoff & Westneat, 2017), but more complex learning reaction norm shapes may require the inclusion of additional non-linear terms in the statistical model (e.g. Moiron, Mathot & Dingemanse, 2018; see Section IV). Therefore, as with reaction norms for non-learning plasticity, 'learning' reaction norms can provide useful statistical descriptions of actual biological processes in the form of data on adjustments in phenotypic values as a result of within- and among-individual differences in learning experiences. If gathered across enough individuals and with information on genetic relatedness, multidimensional learning reaction norm analyses (Fig. 1D) could provide a useful means for estimating genetic and/or permanent environmental effects on learning (e.g. Brandes, 1988; Durisko & Dukas, 2013), including the genetic variance–covariance matrices required to predict evolutionary change in response to selection on learning. This application could prove particularly useful given recent interest in studies of individual differences in learning, innovation and performance in problem-solving tasks [see Carere & Locurto, 2011; Amy, van Oers & Naguib, 2012; Cole & Quinn, 2012; Mathot *et al.*, 2012; Sih & Del Guidice, 2012; Tebbich, Stankewitz & Teschke, 2012; Griffin, Guillette & Healy, 2015; Snell-Rood & Steck, 2019 and other papers in this special issue]. Learning reaction norms thus provide a general foundation for theoretical and statistical conceptualisations of optimal learning strategies and adaptive phenotypic responses to the cumulative experience of a sequence of events.

Figure 1 provides a simplified graphical introduction to the notion of learning reaction norms and their utility, but any workable multidimensional reaction norm framework needs to incorporate learning and forgetting as distinct processes. This is achieved by statistically separating (*i*) the cumulative effect of a sequence of exposure events on learning from (*ii*) the effect of time between these exposure events on forgetting, as represented by two separate *x*-axes in Fig. 2. Such an approach is required whenever the rates of learning and forgetting arise from two distinct processes each with

their own timescale. For instance, in Section I we cite some classic examples involving contrasting timescales of learning *versus* forgetting in different adaptive ecological contexts. Figure 2 illustrates four of these examples, with Fig. 2A,B showing the interactive effects of fast learning and slow forgetting associated with kin discrimination *via* imprinting (Bolhuis, 1991; Holmes & Mateo, 2007) and conditioned taste aversion to potentially toxic foods (Gustavson, 1977; Nicolaus & Nellis, 1987), respectively. These contrast with Fig. 2C,D that show the relatively slow learning and relatively fast forgetting during habituation (see Shettleworth, 2010) or learning for the purposes of adaptive foraging on ephemeral food sources (Hirvonen *et al.*, 1999; Stephens *et al.*, 2007), respectively. More complex scenarios are obviously possible, but our general point here is that representing the potentially separate processes of learning and forgetting in this way facilitates formal comparisons between theoretical expectations (e.g. of the mechanisms involved) and empirical findings (both observational and experimental), and clarifies why 'learning' (and 'forgetting'), in the broadest sense, can usefully be viewed as a particular subset of phenotypic plasticity.

IV. A MATHEMATICAL DESCRIPTION OF LEARNING REACTION NORMS

A powerful advantage of the reaction norm framework is the mathematical and statistical tools, such as mixed-effect models, that are available for estimating key parameters from real data sets. We illustrate this here by modifying such regression models in the form of the 'phenotypic equation' in order to encompass the biological processes involving 'learning' and similar cumulative plasticity effects, but with two important caveats: (*a*) we employ linear equations for convenience by log-transforming what are assumed to be exponential changes in the behavioural response variable (y), although this approach can in most cases accommodate alternative non-linear relationships and more complex mathematical descriptions of the biology (see Section III); and (*b*) we use notation that differs from previous presentations of the statistical parameters in reaction norms (e.g. Nussey *et al.*, 2007) and learning curves from cognitive psychology (see Shettleworth, 2010).

Consider a behavioural outcome y_{ij}, such as the latency to approach a familiar food source during instance i for individual j. We describe this latency as follows:

$$y_{ij} = \beta_0 + I_{0j} + e_{ij}, \tag{1}$$

where β_0 represents the mean population (intercept) value of the behaviour, I_{0j} is the mean deviation from β_0 for individual j, and e_{ij} is the residual unexplained deviation from $\beta_0 + I_{0j}$ in the behaviour at instance i for individual j. A sudden environmental change (e.g. the introduction of a novel object) might alter the behavioural response, but if it is
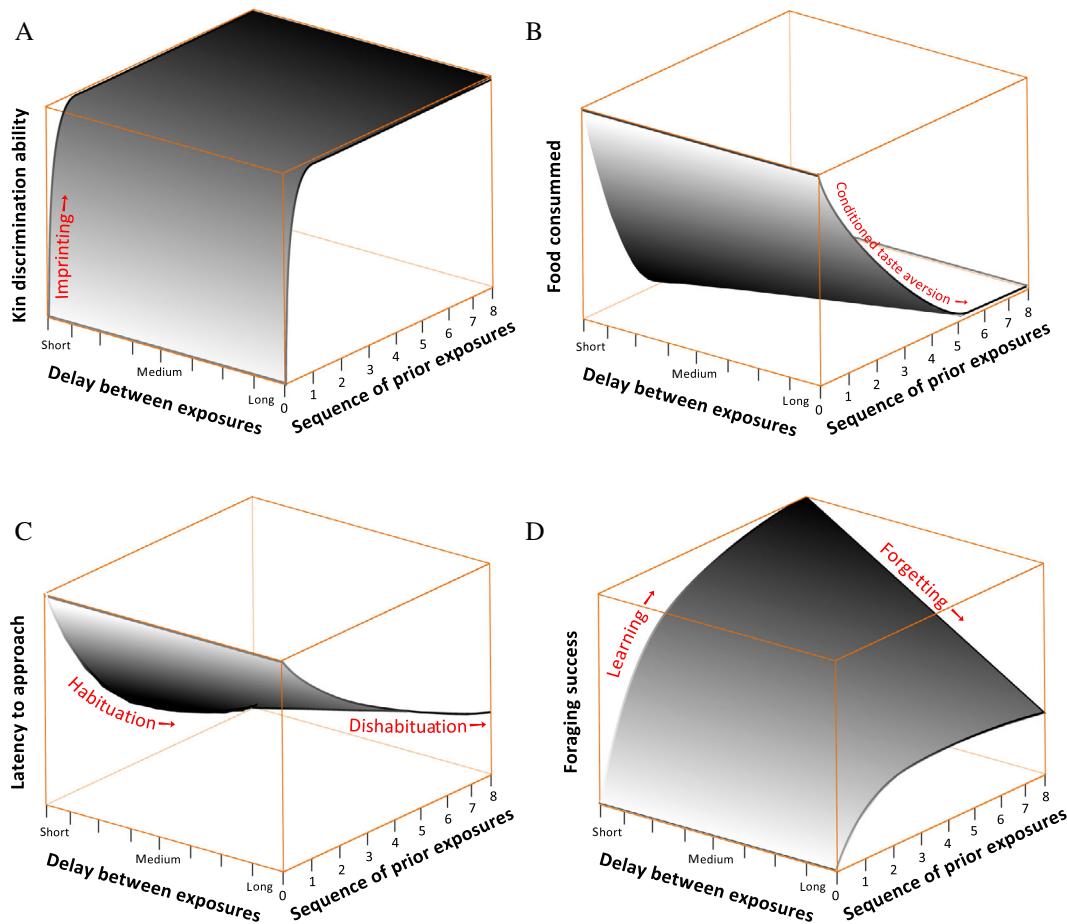
Fig. 2. Illustrations of multidimensional curvilinear learning reaction norms showing different rates of phenotypic plasticity due to the cumulative experience of (*i*) a sequence of prior (reinforcing) exposures *versus* (*ii*) the length of delays between successive exposures in: (A) kin discrimination ability (e.g. in affiliation behaviour towards kin *versus* non-kin) as a result of imprinting requiring usually only one or two prior exposures with little forgetting and hence no effect of time delays; (B) conditioned taste aversion requiring only a small number of prior exposures but with less effect if there are longer delays between those events; (C) habituation to a benign novel object occurring only after a long sequence of similar events and with dishabituation increasing following longer delays between such events; and (D) foraging success increasing *via* slow positive reinforcement (or associative) learning due to experiencing many events in a row with forgetting happening on a similar timescale following increasing delays between events without reinforcement. The darker shading in these grey learning reaction norm surfaces represents the more diminished changes in behaviour in later phenotypic expressions. See main text for details, but note that the particular cases here were chosen for the purposes of illustration. In real systems, these aspects of learning reaction norms can also be linear, and may involve more than just two environmental (*x*-axis) effects.

benign then repeated exposure will often produce habituation (Fig. 1C). This can be done in much the same way as with the phenotypic equation for non-learning behavioural plasticity (see Nussey *et al.*, 2007; Dingemanse *et al.*, 2010; Westneat *et al.*, 2015), but instead of an environmental gradient (E_{ij}) we use a temporal sequence of prior exposures to this new circumstance (X_{ij}) and assay how subjects on average change ($\beta_1 X_{ij}$) their behaviour (i.e. habituate) with each successive experience. We note here that habituation is often an exponential non-linear process over the natural scales of both X and y (see Section II). Thus, phenotypic expressions y_{ij} recorded for each different experience in time allows X_{ij} to capture the magnitude of the time intervals between exposures, and hence it is X_{ij} that can be log-

transformed (e.g. Dingemanse *et al.*, 2012*b*). Alternatively, if the response (y_{ij}) is something like the time an individual takes to achieve some behavioural action (i.e. latency to approach a familiar food source in the presence of a benign novel object) across an ordered number of exposures, then researchers often take account of any non-linearity by log transforming y_{ij} (e.g. Moldoff & Westneat, 2017). The equation then becomes:

$$\log\left(y_{ij}\right) = \beta_0 + I_{0j} + \left(\beta_1 + I_{1j}\right)X_{ij} + e_{ij}. \quad (2)$$

In addition to capturing the hypothesised exponential non-linear relationship, Equation 2 also includes many more

individually unique parameters – not only the estimated intercept at the first exposure to the new circumstance I_{0j}, but also X_{ij} the measured number of prior exposures i per individual j, and the estimated individual slope or change in behaviour per repeated exposure I_{1j} representing individual-specific habituation.

These types of statistical equations usefully reflect specific hypotheses regarding the underlying biology. That is, the parameters that emerge are distillations of underlying biological processes. As mentioned previously, the linear *versus* non-linear nature of habituation is captured by the slope term, and that term should in some way be linked to the cognitive mechanisms of information gathering and storage. Furthermore, the estimated parameters resulting from fitting these equations are expected to fit neatly into theoretical evolutionary processes. For example, the mean rate of habituation to the new context for the population β_1 will evolve if there is selection acting upon among-individual (and presumably among-genotype) variation in this slope. This leads to questions rarely asked about learning: how large is the among-individual variation in habituation in natural populations, how does it arise, and how might selection act on different levels of habituation? While the possibility that there is genetic variation in habituation has been investigated (see Dingemanse *et al.*, 2012*a*,*b*), few of the other questions have been tackled, especially in natural systems. Other non-genetic sources of individual variation should also exist, such as the environmental conditions earlier in the organism's life (e.g. exposure to predation threats during development in stickleback fry; Dingemanse *et al.*, 2012*a*).

Accordingly, we can now assume that our individual subjects j vary in how often they encountered predators during some period preceding our study of habituation (e.g. during early development). We can call this variable X_{2j} and our former variable of number of prior exposures to the benign novel object now becomes X_{1ij}. If we had measured or manipulated X_{2j}, a new equation describes our subjects' responses:

$$\log\left(y_{ij}\right) = \beta_0 + I_{0j} + \left(\beta_1 + I_{1j}\right)X_{1ij} + \beta_2 X_{2j} + \beta_{12}X_{1ij}X_{2j} + e_{ij}.$$

$$(3)$$

This expanded equation captures the types of 'learning' reaction norm surfaces depicted in Fig. 1D or Fig. 2, plus any individual phenotypic variation around those planar shapes, and it can provide a number of insights into the question of how learning might evolve. For instance, if variation in exposure to predators has influenced the responses of our subjects at the population level, then β_2 might explain some of the variance in individual intercepts. In addition, β_{12}, the population-level interaction effect between repeated exposures to X_1 and early-life predator exposure, could explain some of the variance in individual rates of habituation, depending of course upon individual variation in the values of X_1 and X_2 experienced.

This latter term $(\beta_{12}X_{1ij}X_{2j})$ in Equation 3 is particularly interesting, given that learning is a form of plasticity, because it describes nonadditive multidimensional plasticity (Westneat *et al.*, 2019). It is this interaction between learning and non-learning *x*-axes that causes a warped reaction norm surface (e.g. Figs 1D and 2D), and because in this case early-life exposure to predators occurred before the study of habituation, this illustrates a case of a developmental effect on the rate of habituation (see Dingemanse *et al.*, 2010; Stamps & Groothuis, 2010). For simplicity, we assume a similar population-wide effect of predator encounters on all individuals, but there could also be individual variation in the response to X_2. We will skip the full expansion of Equation 3 to include individual (i.e. random) slopes for both X_1 and X_2, plus their interaction, but for more detail on this see Box 2 in Westneat *et al.* (2019). To fit such a model, one would have to measure the effects of exposure to predators (X_{2j}) multiple times per subject (e.g. during development), quantify among-individual variation in these responses (I_{2j}), and conceivably also measure individual variation in the way prior exposure to predators affects habituation (I_{12j}). Such a model would thus make it possible to identify specific parameters relating to individual identity that could be modulating the developmental effects of perceived predation threat on habituation.

With sufficient data and additional information on genetic relatedness, and perhaps also on other traits of interest, existing quantitative genetics approaches could be used to assess underlying genetic variation and covariation in all these reaction norm parameters, and perhaps the selection gradients acting on each of them using multivariate or errors-in-variables models (see Nussey *et al.*, 2007; Dingemanse, Araya-Ajoy & Westneat, 2021). In this way, it is possible to reveal the ecological basis for natural selection acting on the intercepts and the slopes of learning and non-learning reaction norms, and the potential for evolutionary change. We note that recent advances in the statistics of selection allow empirical assessments (i.e. of optimality models of learning) by assessing stabilising selection on reaction norm slopes (see Ponzi *et al.*, 2018; Dingemanse *et al.*, 2021; Martin & Jaeggi, 2022), which opens up additional possibilities for understanding the ecology of selection on learning.

V. CLASSIFYING DIFFERENT FORMS OF LEARNING USING REACTION NORMS

The integration of learning into phenotypic plasticity using the reaction norm framework generates new ideas on the interplay between learning *per se* and conventional non-learning plasticity. For example, learning and other types of plasticity in response to cumulative experience may lead to several types of adjustments within conventional non-learning plasticity reaction norms. Figure 3 shows six different non-mutually exclusive ways in which 'learning' can affect plastic phenotypic responses with potential adaptive consequences.

Note that for reasons of graphical clarity Fig. 3 uses two-dimensional illustrations of how learning might affect linear non-learning reaction norms, when in reality these would constitute multidimensional reaction norms with possible non-linear effects (see Sections II–IV and Equation 3).

(1) Elevation

One impact of learning involves changes to the individual mean phenotype (i.e. Figure 1C) or, in the case of a plastic phenotype, changes to the elevation of the reaction norm in an additive fashion with no effect on the slope (Fig. 3A). Learning that changes the reaction norm elevation might arise *via* cumulative experience from prior events sufficiently closely spaced in time (see Fig. 2), because with sufficient environmental temporal autocorrelation those prior experiences provide useful information regarding an appropriate phenotypic value that adaptively matches the current environment (see Section VIII). It is this learning effect on individual mean trait values that has been explicitly modelled in much of the theoretical literature on learning [see Feldman & Aoki, 2014 and other papers in this special issue; McNamara & Leimar, 2020], but in natural systems we might expect such learning effects on mean individual trait values alongside non-learning plasticity. For example, in passerine birds nestlings beg with greater intensity in response to increasing levels of food deprivation, but within broods the smallest nestlings always beg more than the largest irrespective of food intake rate, and this seems to be due to their learned experience of having to compete more for food within the brood rather than any differences in digestive development, etc. (see Wright *et al.*, 2010*a*; Wetzel *et al.*, 2020).

(2) Slope

Learning could also involve changes in the slope of non-learning reaction norms (Fig. 3B), perhaps due to learned experience of improved pay-offs from modifications in individual responsiveness to variation in the environment. This is the interaction effect between two types of plasticity shown as a warped reaction norm surface in Fig. 1D. For example, individuals might be expected to respond less aggressively when they meet more socially dominant individuals and, because subordinates tend to meet many more dominant individuals, subordinates will benefit more from increased social responsiveness (i.e. by making greater adjustments in their level of aggression based upon the behaviour of their opponents), as compared to dominants that can afford to be more unconditionally aggressive. Accordingly, smaller and/or more subordinate individuals learn with increasing social experience that they need to be more socially responsive (e.g. Koolhaas *et al.*, 1999), and this is represented in Fig. 3B as a non-learning (social) reaction norm that increases in slope with the cumulative effects of learning based upon prior experience. In the more complex case of non-linear reaction norms, then the slope would involve more than just a single (linear) component, and any additional curvilinear components might also be modified separately or in concert

by learning, thereby allowing learning to refine not only the slope but also the 'shape' of the reaction norm.

Many organisms might well exhibit a combination of learned effects on both their reaction norm intercepts *and* slopes at the same time, if only because of the natural covariances that seem to arise between reaction norm elevations and slopes. For example, young stickleback with contrasting early-life experiences regarding the presence of predators tend to show habituation reaction norms to a novel environment that 'fan out' (a positive elevation–slope covariance; Dingemanse *et al.*, 2012*a*). The effect of learning about an increased threat of predation seems both to increase average levels of activity (greater elevation) and to reduce rates of habituation to novel environments (flattening out the negative slope). Therefore, Fig. 3A,B represent relatively simple scenarios for ease of understanding, but mathematical and statistical methods exist to cope with the additional and potentially interesting complexity of most real biological situations (see Section IV).

(3) Environmental estimate error

The memory of prior environmental experiences could affect a non-learning reaction norm by reducing the organism's uncertainty about current environmental conditions, or indicate more clearly where exactly the organism is on the environmental *x*-axis. This has been the topic of extensive theoretical and empirical study, mostly concerning adaptive memory window lengths in the use of past (learnt) information when tracking patch qualities in optimal foraging behaviour and evolutionarily stable learning rules to assess better the phenotypes of competitors or potential social partners, etc. (see Stephens, 1987; Mangel, 1990; Dall, McNamara & Cuthill, 1999; Eliassen *et al.*, 2007; Stephens *et al.*, 2007; Westneat *et al.*, 2015; McNamara & Leimar, 2020). If assessment is not too costly, learning in this context can potentially improve the benefits of plasticity by increasing its precision (i.e. by reducing the 'phenotype–environment mismatch'; Auld, Agrawal & Relyea, 2010), thereby improving the conditions required for the evolution of non-learning plasticity (see Section VIII). This is represented in Fig. 3C as a progressive reduction in environmental error as a result of learning, shown as improvements in the match between estimates of the *perceived* current environment with the true environment on the *x*-axis (i.e. reduced residual variation in the *x*-axis dimension around what is assumed to be the optimal reaction norm). One can imagine a three-dimensional plot of this effect, akin to Fig. 1D, where an absence of learning leads to imprecise assessments of the environment and a cloud of inaccurate phenotypic values surrounding the non-learning reaction norm where the graph indicates zero prior exposures; and where with increasing numbers of prior exposures there is a gradual reduction in the size of this cloud allowing it to resolve into an increasingly precise set of phenotypic expressions that cluster more closely around the reaction norm due to learning. Essentially, information gathered from being repeatedly exposed to a set of cues about the same environmental conditions can be usefully integrated
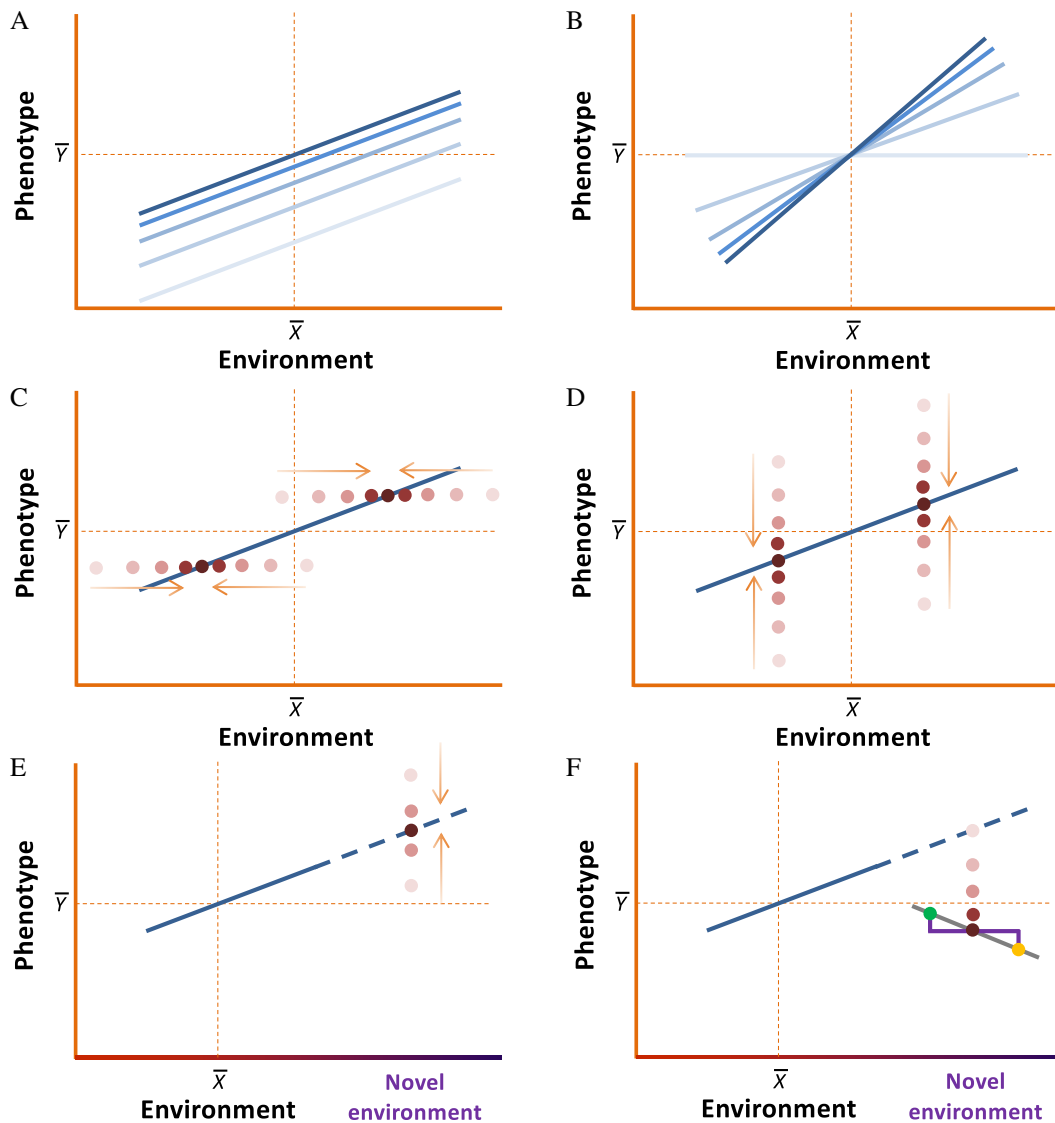
Fig. 3. Conceptual representations of different ways that 'learning' can affect phenotypic values in the context of non-learning reaction norms. For the purposes of illustration, potentially multidimensional reaction norms have been simplified here into two-dimensional representations of linear non-learning reaction norms (in blue), with dots (in red) denoting instances of phenotypic expression. The spacing of the reaction norms and dots indicate the expected non-linear changes over time due to learning from successive prior exposures to the environment allowing the individual to arrive gradually and asymptotically at a new pattern of phenotypic expression, with both reaction norms and dots becoming progressively darker during this process of learning. Learning can affect non-learning reaction norm (A) elevations (mean phenotype) and/or (B) slopes (responsiveness). It can also usefully reduce the degree of error (or residual variation) in instances of phenotypic expression away from optimal reaction norms (illustrated here with just two sets of orange arrows) in either (C) the *x*-dimension, as informational memory from experiencing past environments is used progressively to improve the match between the *perceived* environmental value on the *x*-axis and the true value (see Section VIII), and/or (D) the *y*-dimension, as skills learnt from prior experiences increase the accuracy or precision of the appropriate phenotypic expression given the environment. In addition, we can simplistically represent (E) innovation as an extension of the reaction norm (dashed blue line) in response to novel environmental conditions (in purple) followed by reinforcement learning (based on pay-offs) to refine the expression of a new optimum phenotypic value, and (F) the use of similar innovative learning as a first step in a specific example of reinforcement learning of a novel experimentally imposed optimal reaction norm (grey line or purple dichotomous choice) requiring a series of appropriate learnt behavioural responses *via* training to a particular novel contingency (green *versus* yellow options). See text for further explanation.

into a more precise single estimate of the current environmental conditions *via* some averaging or Bayesian updating procedure within the evolved learning rule

(e.g. McNamara & Houston, 1987). We explore this topic in more detail in Section VIII with a model concerning the evolution of learning, because the accumulation of

informational memory providing more accurate estimates regarding the environment is the other most common type of learning within the theoretical literature on this topic [see Lotem & Halpern, 2012; Feldman & Aoki, 2014 and other papers in this special issue].

(4) Phenotypic precision

Similarly, the reaction norm framework also captures processes that reduce phenotypic error through skills learning (or 'expertise'; see Dukas, 2019), or any other prior experience that improves environmental canalisation or behavioural stability (see Westneat *et al.*, 2015). In this case, learning might reduce the fitness consequences of developmental instability, which has been considered a cost of plasticity (DeWitt *et al.*, 1998; Auld *et al.*, 2010) but is arguably better conceived of as an incomplete benefit from imprecise developmental plasticity (see also Haaland, Wright & Ratikainen, 2020). In Fig. 3D, such learning effects are represented as successive improvements in phenotypic precision towards the individual reaction norm (i.e. reducing residual error variation in the *y*-axis dimension around what is assumed to be an optimal reaction norm). For example, in cooperatively breeding birds helpers-at-the-nest of all ages appear to adjust their level of care in the same way according to current brood demand [i.e. similar reaction norm slopes in what is non-learning reversible plasticity (Wright, 1998; Wright *et al.*, 2010*b*)], but older helpers may benefit from the cumulative experience of prior helping, increasing their skill in finding and gathering appropriate prey items and more correctly matching their levels of help to the needs of young in the nest [i.e. learning (Rowley, 1977; Brown, 1987)]. This too can be visualised as a three-dimensional plot (as in Fig. 1D) with brood demand as the non-learning environmental *x*-axis and level of provisioning experience as the learning prior exposures *x*-axis. We might expect a cloud of imprecise phenotypic expressions (i.e. deviations around the reaction norm line) in the section of the plot indicating no prior provisioning experience, which resolves into a tighter cluster of points closer to the reaction norm surface with better precision in provisioning effort according to brood demand whenever there has been enough prior experience to allow the acquisition of a sufficient level of skill.

The processes depicted in Fig. 3C,D therefore both reduce the size of any residual deviations in phenotypic values (i.e. e_{ij} in Equations 1–3; Section IV), but they do so *via* distinct mechanisms. Hence, they both contribute to reductions in maladaptive biological error in terms of instances of phenotypic expression that deviate from the optimum reaction norm (Westneat *et al.*, 2015, 2019). This assumes that such a reaction norm can be correctly and accurately characterised at the population level, and that any such systematic changes in the residual variation (e_{ij}) can be quantified appropriately in the statistical model (see Westneat *et al.*, 2017). Both processes may also involve feedback from reinforcement learning based upon improved pay-offs from phenotypes more closely approaching optimum reaction norm values, or some other more specific mechanism to improve

accuracy of environmental estimates and/or precision of phenotypic expression. Although all four processes in Fig. 3A–D are unlikely to be mutually exclusive in most biological scenarios (e.g. increases in precision may be accompanied by a directional bias in phenotypic values as well), it is functionally and operationally useful to define separately and quantify statistically each of them. This separation allows us to test properly where and when learning provides adaptive improvements in the position and/or shape of individual multidimensional reaction norms (Fig. 3A,B) and/or in the progressive reduction of residual error in phenotypic expression (Fig. 3C,D).

(5) Innovation

The reaction norm framework we advocate here can also accommodate the many examples of learning in novel and artificial laboratory environments that comprise much of learning research in experimental psychology (see Shettleworth, 2010). These are perhaps best understood initially in the context of innovation learning, which is defined as the adaptive use of new behaviours or phenotypic values in novel contexts (see Reader & Laland, 2003). One of the earliest classic descriptions of innovation was in wild Japanese macaques (*Macaca fuscata*) washing sweet potatoes and floating rice in water to remove the sand mixed in with these novel supplemented food types (Kwai, 1965; for a discussion of this and many other examples, see Reader & Laland, 2003). Innovation trial-and-error learning (Fig. 3E) can be thought of as initially involving the expression of some novel or new level of behaviour based upon applying or extending an existing reaction norm in response to the same or similar cues present in the novel environmental (see Sih, 2013). Innovation success thus depends upon how easily generalised the original environmental cues and experiences are in order to be used effectively by the organism when it finds itself in a particular novel environment with similar or somewhat modified cues (Shettleworth, 2010; Greggor *et al.*, 2019). This process might also often involve multidimensional reaction norms and a combination of different environmental elements (e.g. the stickiness of the sand plus the distance to a water source to wash it off), but for the purposes of illustration here we are restricting our argument to only one environmental *x*-axis. Any extension of existing reaction norm responses for the purposes of innovation might well be followed by a certain amount of skill-type learning (*sensu* Fig. 3D) to improve the precision of the phenotype and thus perhaps also the phenotypic value of the new action (Fig. 3E). Innovation thus also depends upon how easily animals learn and are able to produce the appropriate phenotype once they are faced with a novel environment (see Dukas, 2013), possibly also including reversal learning (see Greggor *et al.*, 2019) when a completely differently shaped adaptive reaction norm is required (see Section V.6 and Fig. 3F). This process usually begins as a result of such learning based upon cumulative experience by one innovating individual, but if successful then it may well spread *via* social learning and the cultural transmission of appropriate

behaviours to other individuals in the social group (see Reader & Laland, 2003).

The process of extending a reaction norm into novel environmental space to produce new phenotypes (Fig. 3E) has been suggested to be a by-product benefit of phenotypic plasticity in general, because it occasionally allows organisms to produce an appropriate phenotype and cope with environmental conditions outside of the normal range for the species (Stephens, 1991; Getty, 1996; West-Eberhard, 2003; Ghalambor *et al.*, 2007; Dayan *et al.*, 2019). Interestingly, rates of innovation correlate positively with relative brain size across species (see Reader & Laland, 2003; Sol *et al.*, 2005; Morand-Ferron, Sol & Lefebvre, 2007), and the ability to innovate effectively by quickly disregarding unsuccessful behavioural options, perhaps without even trying them out, provides an operational description of what many might generally consider 'intelligence' in human and non-human animals (see Ghirlanda, Enquist & Lind, 2014). More anthropomorphic conceptualisations of learning thus focus almost exclusively on trial-and-error reinforcement learning as a general process in isolation, analogous to machine learning and even to natural selection (Watson & Szathmáry, 2016; Watson & Thies, 2019). However, 'learning' in biological organisms will involve a much wider array of processes and effects, and will almost always occur in the context of non-learning plasticity reaction norms (Fig. 3), which then provide the 'prior' phenotype upon which any innovative trial-and-error reinforcement learning must be based.

For empiricists observing attempts at innovation using trial-and-error assessments of different phenotypes, either by simply extending different reaction norms (Fig. 3E) and/or adaptively increasing phenotypic error (i.e. the opposite of the process in Fig. 3D), this can appear as randomly expressed behavioural responses, especially in novel scenarios with no apparent solution. This results in what have been termed 'superstitions' by experimental psychologists, which are idiosyncratic behaviours that individuals acquire for a period of time, such as consistently head bobbing or turning anticlockwise during each repeated attempt to obtain food in an experimental set-up (Skinner, 1948). The maintenance of superstitions over a certain period of time (i.e. in order for the individual to evaluate them thoroughly) can represent an adaptive last resort guess if it involves small costs relative to the potentially large benefits of luckily finding the right answer. This is because, in a stochastic world with no reliable cues to indicate the correct behavioural response such superstitions are better than simply not changing one's phenotype at all or trying completely different phenotypes every time (Foster & Kokko, 2009; Abbott & Sherratt, 2011).

(6) Learning novel contingencies

The same processes involved in innovation learning are likely to be incorporated into the initial phases of learning novel (artificial) contingencies, such as in the experimental protocols used in many studies of learning psychology. Figure 3F illustrates just one example of how in laboratory learning experiments animals might first need to be encouraged to use innovation learning to reproduce suitable (levels of) behaviours based upon some pre-existing reaction norm response, such as by pecking at coloured keys instead of similarly coloured food items when hungry. In this example, the environmental *x*-axis in Fig. 3F could be the key colour (i.e. a spectrum of red to blue) and the *y*-axis indicates the behavioural phenotype of peck rate. In a novel operant set-up, naïve subjects might sometimes obtain food by innovatively pecking a green-coloured key at a high rate whenever present, effectively extending their (blue dashed line) natural reaction norm of pecking more rapidly for seeds in green grass compared to low pecking rates on brown-coloured sand where such food items are fewer and more easily gathered. Over time, skills learning reinforces the response of fewer pecks, since fewer are required to gain food rewards whenever the animal is presented with a coloured-key contingency in the operant set-up (red dots in Fig. 3F). Further reinforcement training could then establish the green/yellow distinction in Fig. 3F – the new (dichotomous purple or linear grey) experimentally imposed reaction norm – by rewarding a slightly higher peck rate when a green-coloured key is present compared to a yellow-coloured key. Only after achieving all this can the specific experiment be carried out to assess individual learning performances on a particular task.

Breaking down the elements of an artificial learning experiment using a reaction norm framework (e.g. Figure 3F) can thus clarify the number of different stages of training needed, depending upon the similarities between the task and the natural environment (and/or previous captive laboratory experiences). It also allows us to appreciate the different processes involved in learning contingencies in artificial contexts, and how they relate to the first four effects of learning we describe in this section (Fig. 3A–D). Identifying these processes may be crucial because artificial learning experiments are used to provide much of the information we have concerning learning mechanisms (see Shettleworth, 2010), and so the various intended or unintended products of any one of the steps involved in the learning of artificial contingencies could affect any conclusions drawn concerning specific learning mechanisms and how they apply to a particular instance of adaptive learning in a natural ecological context.

VI. STATE-DEPENDENT EFFECTS OF ACCUMULATED EXPERIENCE ON PHENOTYPES

A key element of our framework concerns the carryover of cumulative effects from one environmental exposure event to the next within a sequence of individual experiences. The precise mechanisms of this carryover and what mediates and maintains any phenotypic change across episodes of learning (and forgetting) is critical for the expression of learning reaction norms. Here we can utilise the concept of 'state' variables that describe the cumulative (and often non-linear)

increases or decreases in some adaptively important aspect or property of an organism, such as informational state, energy reserves or somatic growth (see Section I). This allows the adaptive nature of carryover effects resulting from accumulated environmental exposures or prior experiences over time to be understood using stochastic dynamic state-dependent models (see Houston & McNamara, 1992, 1999; Clark & Mangel, 2000). Accordingly, these same state-dependent modelling approaches have been used to understand the utility of informational states (e.g. the individual's current estimate or informational memory regarding a foraging environment, Fig. 3C) by applying the same types of Bayesian updating routines as in other mechanistic models of behavioural learning (see Houston & McNamara, 1999; Clark & Mangel, 2000; Dall & Johnstone, 2002; Dunlap & Stephens, 2012). Memory therefore represents just one of a number of different internal neurological, physiological, hormonal and anatomical states that mediate the carryover effects of cumulative environmental experiences and drive 'learning' plasticity (as described above) across different adaptive decision events or time periods in the organism's lifetime (see Section VII).

Learning reaction norms describe changes in phenotypic expression in response to a sequence of cumulative experiences, but part of this process will be captured mid-way through in the phenotypic value of the individual's internal state variable. Therefore, when data are available regarding state variable value(s), this provides a convenient way to split the process of flexible phenotypic expression into two separate steps to facilitate more detailed investigations into the mechanisms behind behavioural and physiological plasticity. The first multidimensional 'learning' reaction norm would plot all of the positive (learning) and negative (forgetting) cumulative environmental effects on the reversibly plastic state variable, as in Fig. 2, but with the internal 'state' variable on the y-axis as an intermediate phenotypic trait. This first learning reaction norm thus summarises any cognitive, physiological or developmental processes affecting the state variable. A second non-learning reaction norm could then subsequently place the state variable itself on the x-axis as an internal environmental variable predicting the phenotypic trait of interest on the y-axis. This second reaction norm would therefore capture any state-dependent phenotypic plasticity in the trait with more direct fitness consequences, rather like a standard (conditional strategy) reaction norm, and without all of the complex multidimensional axes and cumulative environmental effects of the processes involved in Fig. 2.

Separating the processes by which environments affect the state variable from those that translate state into trait expressions has conceptual benefits. For example, the various types and costs of plasticity (DeWitt et al., 1998; Auld et al., 2010) may differ between these two steps, since reliability of informational cues and sampling costs will affect the first step more, but other production costs of plasticity (i.e. those involved in changing the absolute value of expressed trait in that time period and context) may to a greater extent affect

the second step. This decomposition of phenotypic expression into two successive steps is typically more feasible for plasticities involving physiological and morphological states, which can be directly measured in a range of organisms. Assessments of informational states in cognitive studies are likely to be limited mostly to humans, where the reporting of memories or knowledge is somewhat easier. Either way, there are useful parallels here between an individual's accumulated informational state in the cognitive processes behind learning and forgetting (sensu Dall & Johnstone, 2002; Dunlap & Stephens, 2012) and the array of additional physiological and anatomical states that can reflect the cumulative phenotypic effects of development, acclimation and acquisition on the individual as a result of successive environmental experiences. Exploiting these similarities should allow the same theoretical state-dependent modelling techniques and detailed empirical explorations of the processes and quantitative genetics to be used to understand variation within and among individuals in learning plasticity and other analogous biological phenomena.

VII. STUDYING LEARNING *VERSUS* NON-LEARNING PLASTICITY

Given the arguments above, it would seem crucial that researchers are able to decide whether they are studying learning (or other analogous biological phenomena, such as in 'developmental' or 'organisational' plasticity in endocrinology; Nelson & Kriegsfeld, 2017) *versus* other types of non-learning phenotypic plasticity (i.e. 'reversible', 'contextual' or 'activational' plasticity – see Section II). Two practical issues make this important. First, as noted above (Sections II–IV), learning reaction norms would necessarily have an axis capturing the cumulative sequence of prior environmental exposures with the intercept at first exposure, whereas non-learning reaction norms are normally usefully centred on the mean environmental value (or some other suitable metric of centrality) given the assumption that these environmental experiences have independent effects with no obvious temporal ordering. Second, learning *versus* non-learning reaction norms are likely to have distinctly different underlying processes that may suggest different forms of non-linearity, such as the exponential function often expected in 'learning' reaction norms as a result of Bayesian updating and similar learning rules (see Sections II and IV). A key issue in any theoretical or empirical study of phenotypic plasticity is therefore how to decide, either implicitly or explicitly, exactly what to place on the environmental x-axes of any reaction norm, and whether and how to centre x-axes. This will not only define the number and type of explanatory variables in the statistical model, but also the type and duration of the separate phenotypic expressions for each experience or exposure to the environment.

There is an extensive literature on phenotypic plasticity concerned with the types of environmental variables used in

reaction norms, given what we are able to measure empirically or, more interestingly, the environmental cues that organisms themselves are able to perceive either externally or internally, and their different scales and levels of precision (see Moran, 1992; Getty, 1996; Botero *et al.*, 2015; Chevin & Lande, 2015). Adaptive phenotypic plasticity in both learning and non-learning contexts necessarily involves selection for the successful use of perceptual cues that correlate reliably with the environmental variable(s) of importance (Levins, 1963; Lively, 1986; McNamara & Houston, 2009; Fawcett *et al.*, 2012). For example, birds use a variety of cues from day length to temperature and food availability to assess when in the year to start breeding (Brommer, Rattiste & Wilson, 2008; Dawson, 2008; Simmonds, Cole & Sheldon, 2019). Errors made by the organism in such difficult environmental assessments (see Fig. 3C), or imperfect correlations between the cues used and the actual environmental conditions, are important sources of maladaptive residual phenotypic variation from optimum reaction norm values, both within and between individuals (see Westneat *et al.*, 2015, 2017).

By contrast, there has been less explicit discussion about the timescales over which such environmental cues are perceived and utilised in phenotypic plasticity, for example regarding more or less permanent environmental effects and the adaptive nature of memory. Crucially, these timescales are at the heart of an important operational distinction we wish to make here between non-learning phenotypic plasticity in response to current environmental conditions *versus* learning plasticity in response, at least in part, to past information gathered during prior experiences and stored in some internal 'state' variable. From an empirical perspective, it is clear that all phenotypic plasticity takes some sort of minimum duration per 'event'. This is the time that is required for the perception of the informational cue, the cognitive/physiological processing of that information, and its application in the decision to change the level or type of phenotypic expression. Indeed, theoretical treatments of adaptive timescales of phenotypic plasticity necessarily divide time up into discrete decision 'events', which allows us to distinguish between environments perceived and phenotypes expressed in the current 'time event' *versus* in past (and future) time events (see Botero *et al.*, 2015; Tufto, 2015). Logically, in any study the chosen duration of each time event should be determined by the number of identifiably separate decisions involving the type of plasticity of interest that the organism makes relative to its lifespan or some easily defined subset of its lifespan (e.g. a breeding season or life stage). In order to explore the adaptive nature of learning (and forgetting) and other analogous biological phenomena, we need to distinguish between non-learning phenotypic plasticity occurring within the current time event (of a particular length) *versus* learning plasticity that includes the additional cumulative effects of past events on the phenotypic value expressed.

Many studies fail to make a clear distinction regarding the separation of different decision events in time, and this has resulted in some confusion between the terms 'plasticity'

and 'learning', with them often being used almost interchangeably in the behavioural flexibility literature. The evolution of learning then tends to be conflated with the adaptive advantage of the plasticity itself, and so we rarely see specific theoretical or empirical comparisons of the adaptive value of learning (and the additional cognitive costs involved, etc.) over and above normal non-learning plasticity. A good example of this is in game-theoretical modelling of producer *versus* scrounger behaviour in social foraging, described above (see Section III) as a possible empirical example of a phenotype with both learning and non-learning reaction norm *x*-axes (as illustrated by Fig. 1D). In an otherwise quite sophisticated theoretical exploration of the possible frequency-dependent co-existence of fixed *versus* conditional (i.e. plastic) strategies, it is not yet clear under what conditions producer–scrounger conditional strategies should evolve learning plasticity or non-learning plasticity, or both (see Hamblin & Giraldeau, 2009; Dubois *et al.*, 2010; Katsnelson *et al.*, 2012; Afshar & Giraldeau, 2014; Lee *et al.*, 2016; Aplin & Morand-Ferron, 2017). The modelling approach we take in Section VIII might therefore suggest an avenue for future research on this issue.

It is therefore an important decision on the part of the researcher whether they want to explore phenotypic plasticity in response to just the current environmental conditions, or if a more detailed investigative framework is warranted involving plasticity based upon prior exposures and cumulative experiences from past time events (as illustrated in Figs 1–3). This decision involves deciding the most relevant length of a notional 'time event', given the particular study system and research question(s) involved. The example given above of birds using various cues to adjust timing of breeding occurs over a time period encompassing many changing environmental conditions, and so likely involves the use of many possible cumulative experiences, such as changes in day length and phenological changes in food availability (Brommer *et al.*, 2008; Dawson, 2008; Simmonds *et al.*, 2019). Most studies in this area take the expedient short cut of considering each breeding season as a single 'time event' and all cues that occur within it as 'current' to the phenotypic expression of laying date, thereby allowing them to use simpler non-learning reaction norms. We suggest there may be benefits to a more detailed analysis concerning exactly how a particular series of cumulative physical (and social) environmental effects experienced by each individual since arriving on the breeding grounds combine (*via* some Bayesian updating or progressive learning rule) to influence individual laying dates in that season. This would require a more complex phenotypic equation involving an *x*-axis of ordered prior experiences, with multidimensional reaction norms involving 'learning' (see Sections III and IV). Similarly, second-to-second foraging decisions often involve some form of individual (or socially obtained) estimate of prey availability in the current 'time event', such as the expected proportion of successful prey captures within a visit to a foraging patch, and thus optimal foraging studies of sampling effort do not need to include explicit learning processes (see Stephens

et al., 2007; Eliassen *et al.*, 2009; Dall, 2010). However, we could divide the current patch visit into further discrete time divisions and consider each peck or attempted prey capture as a 'time event'. Individual estimates of prey availability would thus be treated as the outcome of an accumulation of remembered experiences from a particular series of successes and (the delays between them as) failures, again summarised by fitting an appropriate Bayesian updating or similar learning rule mechanism (see Fig. 2D). In each case, a more detailed 'learning' plasticity approach has the potential to reveal more complex reaction norms than previously considered, which may or may not be important for our understanding. The question is always: what timescale of discrete 'events' is most effective in capturing the research issue(s) of interest in any study of phenotypic plasticity, and given the appropriate timescale then are there any environmental effects on the phenotype that accumulate across time events, thereby necessitating a 'learning' reaction norm approach as detailed here?

VIII. MODELLING THE EVOLUTION OF LEARNING *VERSUS* NON-LEARNING PLASTICITY

We now present an example to illustrate how one might study the adaptive value of non-learning plasticity *versus* plasticity that incorporates learning and memory. We use the case of a non-learning reaction norm in a varying environment where cues concerning current environmental conditions are not fully reliable (as shown in Fig. 3C). When current cues are only weakly correlated with the fitness-affecting aspect of the environment, phenotypes produced by plastic genotypes may be mismatched to the environment, reducing the benefits of plasticity. Two alternative strategies for improving the effectiveness and thus fitness benefits of plasticity are possible. Primarily, the organism could invest in (*i*) 'sampling' effort, defined here as information gathering to improve the accuracy of any cue or estimate of environmental conditions only during the current 'time event' (*sensu* Stephens, 1987; DeWitt *et al.*, 1998; Dall *et al.*, 1999; Stephens *et al.*, 2007; Eliassen *et al.*, 2007, 2009; Dall, 2010). Alternatively, the organism could also invest in (*ii*) retention and use of an informational memory factor (*sensu* Harley, 1981) regarding environmental cues gathered during past time events. Therefore, 'sampling' is the label we use in this instance (as others have done) for all the effort in information gathering during the current time event (e.g. a visit to a foraging patch), as opposed to 'learning' and use of 'memory' for information retained from sampling during previous time events (e.g. previous visits to the same patch).

Diminishing fitness returns from the accuracy achieved from any information gathering effort (see Hansen, Carter & Pélabon, 2006) means that whenever it is employed then (*i*) sampling will always result in somewhat imperfect information. We are therefore interested in knowing when it might be worth also employing some degree of (*ii*) learning and

informational memory from prior experiences to improve upon such accuracy adaptively, and how this might then in turn also affect the optimal degree of sampling effort and plasticity. Organisms can thus invest more or less time, attention, effort and/or resources into either one of these two options of sampling (i.e. gathering more detailed current information) and learning (i.e. greater use of a memory factor of previously sampled information). However, both come at the cost of such cognitive investment (perception and information processing; Shettleworth, 2010) and/or of other fitness-enhancing activities (i.e. lost opportunity costs; Stephens *et al.*, 2007). So, our model asks the question: when does it pay to evolve learning (and memory) across decision events, as opposed to improving sampling information for plasticity within the current decision event?

Full details of the model are given as online supporting information in Appendix S1, the R code for the model in Appendix S2 and a flow chart of the procedure in Fig. 4, but briefly this individual-based simulation model tracks a population of haploid, asexual individuals in a variable environment. Over many generations, we examine the joint evolution of three unlinked genes coding for (*a*) plasticity (the reaction norm slope), (*b*) a learning memory factor (the weight placed on stored information, using a Rescorla–Wagner rule with exponentially diminishing importance given to more distant past events; see Staddon, 2016), and (*c*) sampling effort (perceptual accuracy invested in the current environmental cue), each with a small mutation probability per generation. The simulations vary in the level of (1) environmental temporal autocorrelation and (2) the reliability of the environmental cue. At each time step or decision event, individuals adjust their phenotype to match the current environment, depending upon their genes for plasticity, the memory factor and sampling effort (see Fig. 4 for a simplified diagram). We assume diminishing returns on sampling (Fig. S1), and for simplicity linearly increasing costs of greater sampling, memory and reaction norm slopes (although qualitatively similar results were obtained using curvilinear costs). Individual fitness (relative lifetime reproductive success) here relates to how well, on average, the individual's phenotype matched the environmental conditions across time steps during their lifetime. In Fig. 5 we describe the conditions promoting the evolution of learning when using exponentially less memory information from each successive step back in time, as illustrated in the left-centred learning reaction norm 'number of prior exposures' *x*-axes in Figs 1 and 2.

As Fig. 5 shows, plasticity only evolves (i.e. reaction norm slope $g_p > 0$, Fig. 5A) if the environmental cues are sufficiently reliable (see Botero *et al.*, 2015; Tufto, 2015). Sampling effort g_s (Fig. 5C) closely tracks the pattern in reaction norm slopes (Fig. 5A), because sampling provides the informational cues necessary for plasticity [i.e. it adaptively reduces the estimation error in the *x*-axis dimension of the reaction norm sufficiently to allow adaptive plasticity to evolve (see Hansen *et al.*, 2006; Westneat *et al.*, 2019)]. As with other costs of plasticity, this effectively means that it is the costs of sampling
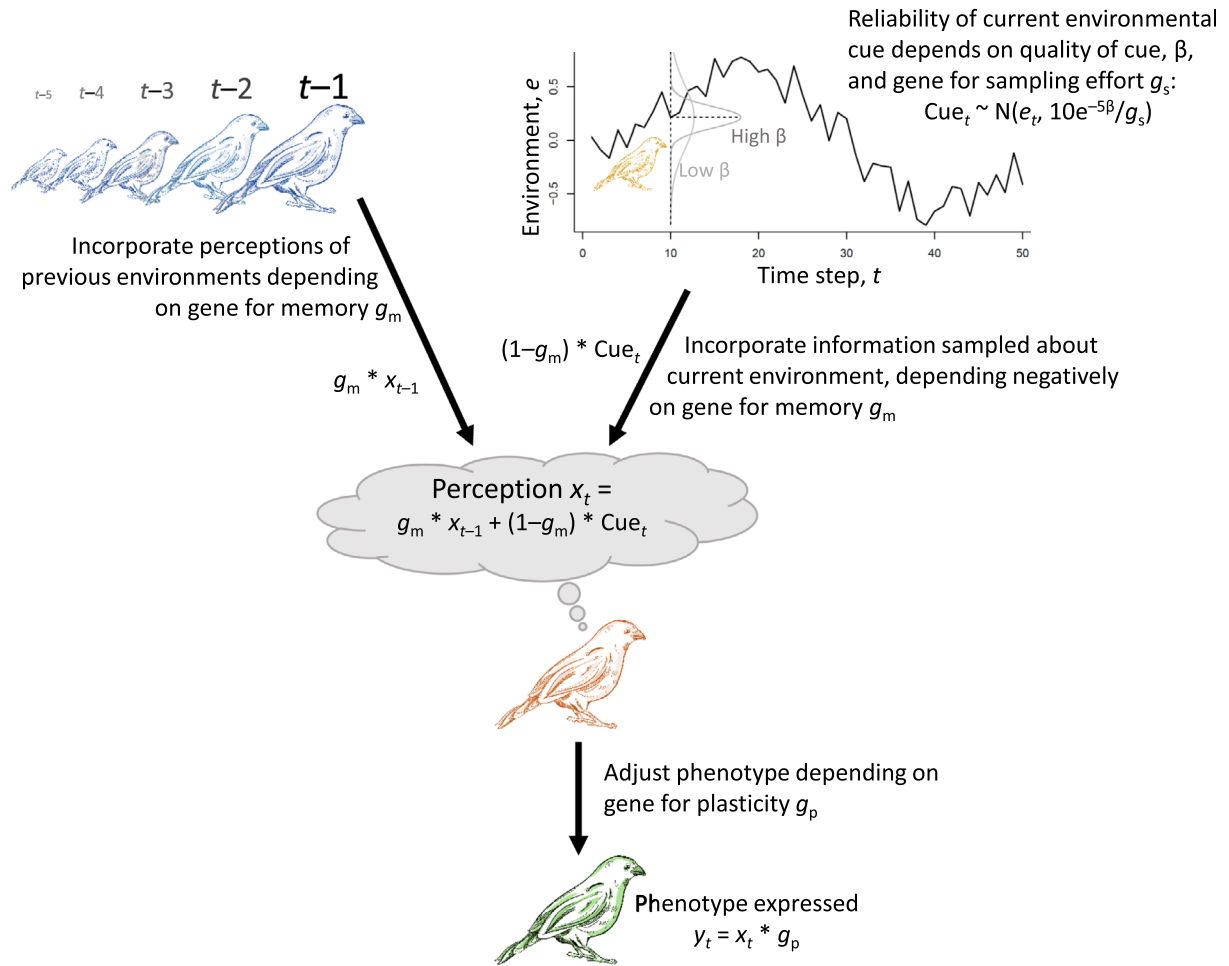
Fig. 4. Flowchart of the individual-based simulation model procedure for phenotypic plasticity. The individual's current perception of its environment x_t (grey thought-bubble) depends upon both its memory of past cues (left path) and the information it sampled during the current time step Cue_t (right path), and their relative weight (g_m and $1-g_m$ respectively). The phenotype y_t is then plastically adjusted towards a match with the current perceived environment. See Appendix S1 for more details. Haploid bird illustration: Wikimedia commons.

(DeWitt *et al.*, 1998) that limit the adaptive upper limit of reaction norm slopes in most cases to less than their maximal value of 1 (compare the upper *versus* lower panels, low *versus* high sampling costs, in Fig. 5A). However, unlike the reaction norm slopes, sampling effort drops away again at higher values of environmental cue reliability (upper panel Fig. 5C), because less investment in sampling is needed to achieve the same levels of informational accuracy concerning current environmental conditions.

As might be expected, learning and use of the memory factor g_m only evolves when there are high levels of environmental temporal autocorrelation (Fig. 5B). Notably, the highest values of the memory factor evolve at lower levels of environmental cue reliability, as compared with the peak in evolution of sampling effort. At intermediate cue reliabilities (e.g. $\beta = 0.5$), plasticity can only evolve when the temporal autocorrelation in environments is sufficiently high (i.e. $\beta > 0.5$), because it is this that allows learning to evolve

and provide the reliable information needed for plasticity to evolve (Fig. 5A). This suggests that memory of prior environmental conditions allows greater levels of plasticity (i.e. steeper reaction norms slopes) to be adaptive when cues concerning current environmental conditions are less reliable and sampling itself is too costly (compare upper *versus* lower panels Fig. 5A,B). An example of this might be birds that face increasingly unpredictable weather as winter approaches only being able to make adaptive daily adjustments in their fat stores by using memories of weather conditions the day before as a guide because this provides the only sufficiently accurate predictor of current daily temperature variation (Ratikainen & Wright, 2013).

A more complete example of such phenomena is provided by the effects of temporal autocorrelations in brood demand on parental provisioning effort in birds. Parents often use information going back in time up to two previous nest visits in order to adjust their level of nestling care more accurately
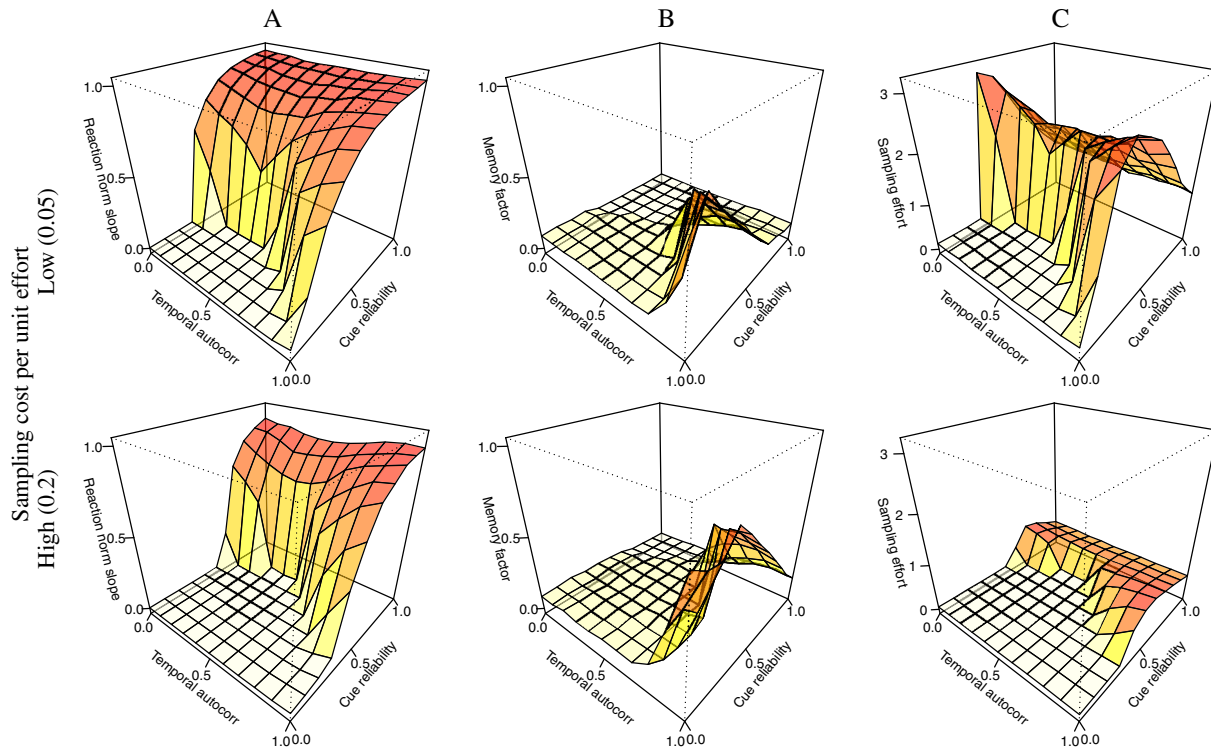
Fig. 5. Individual-based simulation model results, showing evolved genetic values for: (A) plasticity or the slope of the reaction norm g_p ($g_p = 1$ maximises phenotype–environment matching); (B) investment in learning in terms of a memory factor g_m for the use of knowledge regarding environmental conditions during previous decision events; and (C) sampling effort during the current decision or time event, g_s. Results are given according to variation in the reliability of environmental cues β (i.e. how correlated they are with the fitness-impacting environmental factor), and how temporally autocorrelated the environmental factor itself is from one decision event to the next, α. Simulations involved a population of 200 individuals with 50 decision events or time steps per lifetime. Results are shown after 1000 generations averaged across 20 replicates per grid square. In the top panel in each case, the cost of plasticity = 0.05 per unit of reaction norm slope; cost of memory = 0.05 times the proportional use of the memory factor; and cost of sampling = 0.05 per unit sampling effort. In the bottom panel, the costs of plasticity and memory are the same, but with an increased cost of sampling = 0.2. See main text for explanation and Appendix S1 for more details on the model.

(Westneat *et al.*, 2017), presumably because more accurate lengthy (sampling) assessments of brood demand during just one extended visit are too costly for both parents and nestlings. Interestingly, provisioning by additional helpers-at-the-nest in cooperative groups decreases the temporal autocorrelations in brood demand experienced by individuals over successive visits to the nest, resulting in only begging experienced during the very last visit being used to adjust provisioning effort in these systems (Wright *et al.*, 2010*b*). This lack of temporal autocorrelation in brood need in cooperative breeders and the lower reliability of remembered information concerning prior begging events also causes a greater number of last-minute adjustments to current brood begging upon arrival back at the nest in terms of the amount of food in the bill that is actually then fed to the nestlings (Wright, 1997; McDonald, Kazem & Wright, 2007).

The model presented here provides a general example of the value of explicit definitions regarding 'time steps' and 'decision events' when distinguishing non-learning plasticity (the use of sampling) from learning plasticity (the additional use of a memory factor) – see Section VII. More importantly,

it shows how temporal autocorrelations in the environment are a key requirement for the adaptive evolution of any type of learning effect within phenotypic plasticity. This requirement for environmental temporal autocorrelation in order for learning to evolve, along with recent models showing the evolution of plasticity only in the context of variable and predictable-enough environments (see Botero *et al.*, 2015; Tufto, 2015), would appear to resolve earlier debates concerning whether more or less variable environmental regimes promote the evolution of learning (see Stephens, 1991, 1993; Papaj, 1994; Kerr & Feldman, 2003). This is supported in an elegant series of selection experiments on oviposition decisions in *Drosophila* (Dunlap & Stephens, 2009), which demonstrate that evolution of learning depends upon reliable-enough cues in a sufficiently temporally autocorrelated environment, similar to our results in Fig. 5. Together, environmental predictability *via* reliable cues and sufficient temporal autocorrelation in environmental variation are therefore expected to determine the conditions for the adaptive evolution of learning and analogous biological phenomena.

The evolution of 'learning' could also be usefully modelled in the context of the other types of modifications to non-learning plasticity reaction norms, as described in Section V. So, although in this model we only explore the adaptive value of a multidimensional reaction norm that includes prior experience *versus* a unidimensional norm that includes only sampling of the environmental *x*-axis (i.e. Figure 3C), the technical approach taken should be illustrative of the ways that other such models could be generated (and even empirical assessments made) for learning adjustments of reaction norm elevations, slopes and phenotypic precision in plasticity (Fig. 3A–D). It is also important to remember that any interaction between learning and non-learning reaction norms has the potential to work in both directions. Hence, non-learning plasticity can equally affect the elevations, slopes, accuracies and precisions involved in learning reaction norms. For example, following the discussion in Section IV, plastic responses in stickleback fry to variation in the threat of predation influences individual rates of habituation (Dingemanse *et al.*, 2012a). More complex extensions to these models could also explicitly explore the evolution of multidimensional learning reaction norms by including more than one temporally ordered *x*-axis, allowing the evolution of various combinations of positive 'learning' effects due to cumulative recent experiences *versus* negative 'forgetting' effects due to time delays between those experiences, as illustrated in Fig. 2. The advantage of such models over previous studies comparing the relative performance of specifically chosen learning rules is that, given sufficiently flexible mathematical representations, the optimum reaction norm in each case would be allowed to evolve freely to a set of genuine evolutionarily stable values (for a useful discussion of this point, see Fawcett *et al.*, 2012). In this way, it becomes possible to explore systematically exactly why different adaptive contexts result in contrasting shapes of learning reaction norms (as in Fig. 2), and whether these can all be described by simple changes in key parameter values in a general learning rule or whether specifically different learning rule equations are needed for very different biological contexts. Indeed, as with the model results presented here (Fig. 5), one major goal of such theoretical explorations would be to identify the conditions under which no learning evolves.

Additional modelling approaches can also help researchers explore the full range of effects of environmental cue predictability and temporal autocorrelation on the evolution of learning (and other analogous biological phenomena) across successive decision events within a lifetime. As argued above (Section VI), state-dependent stochastic dynamic programming would appear to be the most appropriate method, since it has already been used in the context of sampling and learning in foraging (Houston & McNamara, 1999; Clark & Mangel, 2000; Dall & Johnstone, 2002; Dunlap & Stephens, 2012). Such models can usefully integrate optimal learning rules with plasticity evolving under more or less temporally autocorrelated and stochastic or uncertain environmental conditions, thereby testing the full range of adaptive possibilities for learning

and non-learning reaction norms, as well as their performance under different levels of environmental predictability *versus* stochasticity. We therefore already appear to have the theoretical tools available to assess critically and understand learning and memory and other analogous biological phenomena in phenotypic plasticity from a reaction norm perspective. This approach could also be used to assess formally the role of plasticity and especially learning in ensuring population viability in changing environments due to rapid anthropogenic change (see Greggor *et al.*, 2019).

IX. CONCLUSIONS

(1) Learning (and forgetting) are forms of phenotypic plasticity that involve the influences of cumulative prior experience(s) of an environmental cue on the value of a trait. By utilising reaction norms with an environmental axis that captures the sequence of environmental exposures (and possibly their timing), learning along with other similar biological phenomena (e.g. development, acquisition, acclimation) can thus be suitably accommodated within existing theoretical and empirical approaches to phenotypic plasticity, with implications for our understanding of the evolutionary ecology of learning and of phenotypic plasticity more generally.
(2) A reaction norm approach provides direct links with statistical methods that suitably capture the structure of data collected on labile phenotypes, and the parameters that feed into evolutionary analyses, and thus with biological hypotheses regarding phenotypic plasticity. In the case of learning (and analogous biological phenomena), the reaction norm approach is modified by defining environmental *x*-axes that represent an individual's cumulative sequence of environmental exposures or experiences. Doing so clarifies ongoing methodological issues within wider studies of phenotypic plasticity, such as non-linear and multidimensional reaction norms, and the consequences of locating intercepts along such environmental *x*-axes. Moreover, integrating learning within the broader framework of phenotypic plasticity provides a robust theoretical and conceptual framework for understanding the diversity of forms of learning and how learning *per se* may interact with conventional types of non-learning plasticity.
(3) Learning reaction norms based upon cumulative prior experience are conceptually familiar as habituation and learning curves in psychology, and as the outcome of learning rules in evolutionary models of learning. However, a reaction norm framework allows us to generate new hypotheses about how cumulative experience interacts with conventional plasticity. Prior experience can alter (*a*) the mean phenotype [elevation] and (*b*) responsiveness [slope] to another environmental factor, or could reduce error in either (*c*) environmental estimates [*via* informational memory] or (*d*) consistency in phenotypic precision [*via* skill acquisition]. Our approach can also encompass (*e*) innovation and (*f*) learning of novel contingencies in artificial (laboratory) settings, generating new insights into studies

of learning mechanisms in captivity and how they might relate to adaptive phenotypic plasticity in an ecological context.

(4) An important research question is whether or not learning or analogous biological processes are involved in a particular example of phenotypic plasticity. Answering this requires the use of a suitable definition for separate time (or decision) 'events' in any particular example of phenotypic plasticity under study, as well as the potential for state variables that allow organisms to track variation in temporally autocorrelated environments.

(5) Our individual-based simulation model demonstrates that memory state concerning previous environmental conditions may be more effective at promoting the evolution of plasticity than additional investment in sampling information about the current environment, favouring the evolution of multidimensional plasticity involving past experience and current environmental conditions. Many natural environments are temporally autocorrelated, and so 'learning' (in its most general sense) may often represent an overlooked component of phenotypic plasticity in many cases where studies failed to distinguish between learning *versus* non-learning reaction norms.

(6) Opportunities exist to apply this reaction norm approach to the evolution of 'learning' across a wide array of fields and natural ecological contexts, and to connect these with laboratory studies concerning learning mechanisms and cognition. This would also allow us to test formally the behavioural gambit (the assumption that psychological mechanisms do not constrain the expression of adaptive behaviour; Fawcett *et al.*, 2012) as it pertains to learning as plasticity, because it has been suggested that there are constraints on the evolution of different learning mechanisms due to their utilisation in common cognitive functions (Lotem, 2013). Such constraints would then need to be factored into any reaction norm approach to learning. Therefore, our understanding of the evolution of some of the more complex and interesting forms of plasticity may be advanced more rapidly and effectively by applying a learning reaction norm approach.

X. ACKNOWLEDGEMENTS

XI. REFERENCES

Abbott, K. R. & Sherratt, T. N. (2011). The evolution of superstition through optimal use of incomplete information. *Animal Behaviour* **82**, 85–92.

Afshar, M. & Giraldeau, L.-A. (2014). A unified modelling approach for producer-scrounger games in complex ecological conditions. *Animal Behaviour* **96**, 167–176.

Agrawal, A. A. (2001). Phenotypic plasticity in the interactions and evolution of species. *Science* **294**, 321–326.

Amy, M., van Oers, K. & Naguib, M. (2012). Worms under cover: relationships between performance in learning tasks and personality in great tits (*Parus major*). *Animal Cognition* **15**, 763–770.

Angilletta, M. J., Niewiarowski, P. H. & Navas, C. A. (2002). The evolution of thermal physiology in ectotherms. *Journal of Thermal Biology* **27**, 249–268.

Aoki, K. & Feldman, M. W. (2014). Evolution of learning strategies in temporally and spatially variable environments. *Theoretical Population Biology* **91**, 3–19.

Aplin, L. M. & Morand-Ferron, J. (2017). Stable producer-scrounger dynamics in wild birds: sociability and learning speed covary with scrounging behaviour. *Proceedings of the Royal Society London B* **284**, 20162872.

Arnold, P. A., Kruuk, L. E. B. & Nicotra, P. A. (2019). How to analyse plant phenotypic plasticity in response to a changing climate. *New Phytologist* **222**, 1235–1241.

Auld, J. R., Agrawal, A. A. & Relyea, R. A. (2010). Re-evaluating the costs and limits of adaptive phenotypic plasticity. *Proceedings of the Royal Society London B* **277**, 503–511.

Balda, R. P., Kamil, A. C. & Bednekoff, P. A. (1996). Predicting cognitive capacity from natural history. In *Current Ornithology* (Volume **13**, eds V. N. Nolan and E. D. Ketterson), pp. 33–66. Plenum Press, New York.

Barnard, C. J. & Sibly, R. M. (1981). Producers and scroungers: a general model and its application to captive flocks of house sparrows. *Animal Behaviour* **29**, 543–550.

Beaman, J. E., White, C. R. & Seebacher, F. (2016). Evolution of plasticity: mechanistic links between development and reversible acclimation. *Trends in Ecology and Evolution* **31**, 237–249.

Becker, J. B., Breedlove, S. M., Crews, D. & McCarthy, M. M. (2002). *Behavioral Endocrinology*, Second Edition. MIT Press, Cambridge.

Belmaker, A., Motro, U., Feldman, M. W. & Lotem, A. (2012). Learning to choose among social foraging strategies in adult house sparrows (*Passer domesticus*). *Ethology* **118**, 1111–1121.

Bernays, E. A. & Chapman, R. E. (1994). *Host-Plant Selection by Phytophagous Insects*. Chapman & Hall, New York.

Bills, A. G. (1934). *General Experimental Psychology*. Longmans Psychology Series, pp. 192–215. Longmans, Green & Co, New York.

Bolhuis, J. J. (1991). Mechanisms of avian imprinting: a review. *Biological Reviews* **66**, 303–345.

Botero, C. A., Weissing, F. J., Wright, J. & Rubenstein, D. R. (2015). Evolutionary tipping points in the capacity to adapt to environmental change. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 184–189.

Brandes, C. (1988). Estimation of learning behavior in honeybees (*Apis mellifera capensis*). *Behavior Genetics* **18**, 119–132.

Brommer, J. E., Rattiste, K. & Wilson, A. J. (2008). Exploring plasticity in the wild: laying date–temperature reaction norms in the common gull *Larus canus*. *Proceedings of the Royal Society London B* **275**, 687–693.

Brown, J. L. (1987). *Helping and Communal Breeding in Birds: Ecology and Evolution*. Princeton University Press, Princeton.

Brown, R. L. (2013). Learning, evolvability and exploratory behaviour: extending the evolutionary reach of learning. *Biology and Philosophy* **28**, 933–955.

Carere, C. & Locurto, C. (2011). Interaction between animal personality and animal cognition. *Current Zoology* **57**, 491–498.

Catchpole, C. & Slater, P. (1995). *Bird Song: Biological Themes and Variations*. Cambridge University Press, Cambridge.

Chevin, L. M. & Lande, R. (2015). Evolution of environmental cues for phenotypic plasticity. *Evolution* **69**, 2767–2775.

Clark, C. W. & Mangel, M. (2000). *Dynamic State Variable Models in Ecology: Methods and Applications*. Oxford University Press, Oxford.

Cole, E. F. & Quinn, J. L. (2012). Personality and problem-solving performance explain competitive ability in the wild. *Proceedings of the Royal Society London B* **279**, 1168–1175.

Dall, S. R. X. (2010). Managing risk: the perils of uncertainty. In *Evolutionary Behavioral Ecology* (eds D. F. Westneat and C. W. Fox), pp. 194–206. Oxford University Press, Oxford.

Dall, S. R. X., Giraldeau, L.-A., Olsson, O., McNamara, J. M. & Stephens, D. W. (2005). Information and its use by animals in evolutionary ecology. *Trends in Ecology and Evolution* **20**, 187–193.

Dall, S. R. X. & Johnstone, R. A. (2002). Managing uncertainty: information and insurance under the risk of starvation. *Philosophical Transactions of the Royal Society London B* **357**, 1519–1526.

Dall, S. R. X., McNamara, J. M. & Cuthill, I. C. (1999). Interruptions to foraging and learning in a changing environment. *Animal Behaviour* **57**, 233–241.

DAWSON, A. (2008). Control of the annual cycle in birds: endocrine constraints and plasticity in response to ecological variability. *Philosophical Transactions of the Royal Society London B* **1497**, 1621–1633.

DAYAN, D. I., GRAHAM, M. A., JOHN, G., BAKER, J. A. & FOSTER, S. (2019). Incorporating the environmentally sensitive phenotype into evolutionary thinking. In *Evolutionary Causation: Biological and Philosophical Reflections* (eds T. ULLER and K. N. LALAND), pp. 81–107. MIT Press, London.

DEWITT, T. J., SIH, A. & WILSON, D. S. (1998). Cost and limits of phenotypic plasticity. *Trends in Ecology and Evolution* **13**, 77–81.

DINGEMANSE, N. J., ARAYA-AJOY, Y. G. & WESTNEAT, D. F. (2021). Most published selection gradients are underestimated: why this is and how to fix it. *Evolution* **75**, 806–818.

DINGEMANSE, N. J., BARBER, I., WRIGHT, J. & BROMMER, J. (2012*a*). Quantitative genetics of behavioural reaction norms: genetic correlations between personality and behavioural plasticity vary across stickleback populations. *Journal of Evolutionary Biology* **25**, 485–496.

DINGEMANSE, N. J., BOUWMAN, K. M., VAN DE POL, M., VAN OVERVELD, T., PATRICK, S. C., MATTHYSEN, E. & QUINN, J. L. (2012*b*). Variation in personality and behavioural plasticity across four populations of the great tit *Parus major*. *Journal of Animal Ecology* **81**, 116–126.

DINGEMANSE, N. J., KAZEM, A. J. N., RÉALE, D. & WRIGHT, J. (2010). Behavioural reaction norms: animal personality meets individual plasticity. *Trends in Ecology and Evolution* **25**, 81–89.

DINGEMANSE, N. J. & WOLF, M. (2013). Between-individual differences in behavioural plasticity: causes and consequences. *Animal Behaviour* **85**, 1031–1039.

DUBOIS, F., MORAND-FERRON, J. & GIRALDEAU, L.-A. (2010). Learning in a game context: strategy choice by some keeps learning from evolving in others. *Proceedings of the Royal Society London B* **277**, 3609–3616.

DUKAS, R. (1998). Evolutionary ecology of learning. In *Cognitive Ecology* (ed. R. DUKAS), pp. 129–174. University of Chicago Press, Chicago.

DUKAS, R. (2004). Evolutionary biology of animal cognition. *Annual Review of Ecology, Evolution, and Systematics* **35**, 347–374.

DUKAS, R. (2013). Effects of learning on evolution: robustness, innovation and speciation. *Animal Behaviour* **85**, 1023–1030.

DUKAS, R. (2017). Cognitive innovations and the evolutionary biology of expertise. *Philosophical Transactions of the Royal Society London B* **372**, 20160427.

DUKAS, R. (2019). Animal expertise: mechanisms, ecology and evolution. *Animal Behaviour* **147**, 199–210.

DUKAS, R., CLARK, C. W. & ABBOTT, K. (2006). Courtship strategies of male insects: when is learning advantageous? *Animal Behaviour* **72**, 1395–1404.

DUNLAP, A. S., CHEN, B. B., BEDNEKOFF, P. A., GREENE, T. G. & BALDA, R. P. (2006). A state-dependent sex difference in spatial memory in pinyon jays, *Gymorhinus cyanocephalus*: mated females forget as predicted by natural history. *Animal Behaviour* **72**, 401–411.

DUNLAP, A. S. & STEPHENS, D. W. (2009). Components of change in the evolution of learning and unlearned preferences. *Proceedings of the Royal Society London B* **267**, 3201–3208.

DUNLAP, A. S. & STEPHENS, D. W. (2012). Tracking a changing environment: optimal sampling, adaptive memory and overnight effects. *Behavioural Processes* **89**, 86–94.

DURISKO, Z. & DUKAS, R. (2013). Effects of early-life experiences on learning ability in fruit flies. *Ethology* **119**, 1067–1076.

ELIASSEN, S., JORGENSEN, C., MANGEL, M. & GISKE, J. (2007). Exploration or exploitation: life expectancy changes the value of learning in foraging strategies. *Oikos* **116**, 513–523.

ELIASSEN, S., JØRGENSEN, C., MANGEL, M. & GISKE, J. (2009). Quantifying the adaptive value of learning in foraging behaviour. *American Naturalist* **174**, 478–489.

ENSMINGER, A. L. & WESTNEAT, D. F. (2012). Individual and sex differences in habituation and neophobia in house sparrows (*Passer domesticus*). *Ethology* **118**, 1085–1095.

FAWCETT, T. W., HAMBLIN, S. & GIRALDEAU, L.-A. (2012). Exposing the behavioral gambit: the evolution of learning and decision rules. *Behavioral Ecology* **24**, 2–11.

FELDMAN, M. W. & AOKI, K. (2014). Preface to the theoretical population biology special issue on learning. *Theoretical Population Biology* **91**, 1–2.

FLAJNIK, M. F. & KASAHARA, M. (2010). Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nature Reviews Genetics* **11**, 47–59.

FOSTER, K. R. & KOKKO, H. (2009). The evolution of superstitious and superstitious-like behaviour. *Proceedings of the Royal Society London B* **276**, 31–37.

GAVRILETS, S. & SCHEINER, S. M. (1993). The genetics of phenotypic plasticity. V. Evolution of reaction norm shape. *Journal of Evolutionary Biology* **6**, 31–48.

GETTY, T. (1996). The maintenance of phenotypic plasticity as a signal detection problem. *American Naturalist* **148**, 378–385.

GHALAMBOR, C. K., MCKAY, J. K., CARROLL, S. P. & REZNICK, D. N. (2007). Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology* **21**, 394–407.

GHIRLANDA, S., ENQUIST, M. & LIND, J. (2014). Coevolution of intelligence, behavioral repertoire, and lifespan. *Theoretical Population Biology* **91**, 44–49.

GRAFEN, A. (1984). Natural selection, kin selection and group selection. In *Behavioural Ecology: An Evolutionary Approach*, Second Edition (eds J. R. KREBS and N. B. DAVIES), pp. 62–84. Blackwell Scientific, Oxford.

GREGGOR, A. L., TRIMMER, P. C., BARRETT, B. J. & SIH, A. (2019). Challenges of learning to escape evolutionary traps. *Frontiers in Ecology and Evolution* **7**, 408.

GRIFFIN, A. S., GUILLETTE, L. M. & HEALY, S. D. (2015). Cognition and personality: an analysis of an emerging field. *Trends in Ecology and Evolution* **30**, 207–214.

GUSTAVSON, C. R. (1977). Comparative and field aspects of learned food aversions. In *Learning Mechanisms in Food Selection* (eds L. M. BARKER, M. R. BEST and M. DOMJAN). Baylor University Press, Waco.

HAALAND, T. R., WRIGHT, J. & RATIKAINEN, I. I. (2020). Generalists versus specialists in fluctuating environments: a bet-hedging perspective. *Oikos* **129**, 879–890.

HAMBLIN, S. & GIRALDEAU, L.-A. (2009). Finding the evolutionarily stable learning rule for frequency-dependent foraging. *Animal Behaviour* **78**, 1343–1350.

HANSEN, T. F., CARTER, A. J. R. & PÉLABON, C. (2006). On adaptive accuracy and precision in natural populations. *American Naturalist* **168**, 168–181.

HARLEY, C. B. (1981). Learning the evolutionarily stable strategy. *Journal of Theoretical Biology* **89**, 611–633.

HEALY, S. D. & HURLY, T. A. (2004). Spatial learning and memory in birds. *Brain, Behavior and Evolution* **63**, 211–220.

HIRVONEN, H., RANTA, E., RITA, H. & PEUHKURI, N. (1999). Significance of memory properties in prey choice decisions. *Ecological Modelling* **115**, 177–189.

HOLLIS, K. L., PHARR, V. L., DUMAS, M. J., BRITTON, G. B. & FIELD, J. (1997). Classical conditioning provides paternity advantage for territorial male blue gouramis (*Trichogaster trichopterus*). *Journal of Comparative Psychology* **111**, 219–225.

HOLMES, W. G. & MATEO, J. M. (2007). Kin recognition in rodents: issues and evidence. In *Rodent Societies: An Ecological and Evolutionary Perspectives* (eds J. O. WOLFF and P. W. SHERMAN), pp. 216–228. Chicago University Press, Chicago.

HOUSTON, A. I. & MCNAMARA, J. M. (1992). Phenotypic plasticity as a state-dependent life-history decision. *Evolutionary Ecology* **6**, 243–253.

HOUSTON, A. I. & MCNAMARA, J. M. (1999). *Models of Adaptive Behaviour: An Approach Based on State*. Cambridge University Press, Cambridge.

HSU, Y., RYAN, R. L. & WOLF, L. L. (2006). Modulation of aggressive behaviour by fighting experience: mechanisms and contest outcomes. *Biological Reviews* **81**, 33–74.

JABER, M. Y. (ed.) (2011). *Learning Curves: Theory, Models and Applications*. CRC Press, Taylor & Francis Group, Boca Raton.

JACKSON, R. R. & CARTER, C. M. (2001). Geographic variation in reliance on trial-and-error signal derivation by *Portia labiata*, an araneophagic jumping spider from the Philippines. *Journal of Insect Behavior* **14**, 799–827.

JANEWAY, C. A., TRAVERS, P., WALPORT, M. & SHLOMCHIK, M. J. (2005). *Immunobiology*, Sixth Edition. Garland Science, New York & London.

JOHNSTON, T. D. (1982). Selective costs and benefits in the evolution of learning. *Advances in the Study of Behavior* **12**, 65–106.

KATSNELSON, E., MOTRO, U., FELDMAN, M. W. & LOTEM, A. (2008). Early experience affects producer–scrounger foraging tendencies in the house sparrow. *Animal Behaviour* **75**, 1465–1472.

KATSNELSON, E., MOTRO, U., FELDMAN, M. W. & LOTEM, A. (2012). Evolution of learned strategy choice in a frequency-dependent game. *Proceedings of the Royal Society London B* **279**, 1176–1184.

KAWECKI, T. J. & STEARNS, S. C. (1993). The evolution of life histories in spatially heterogeneous environments: optimal reaction norms revisited. *Evolutionary Ecology* **7**, 155–174.

KERR, B. & FELDMAN, M. W. (2003). Carving the cognitive niche: optimal learning strategies in homogeneous and heterogeneous environments. *Journal of Theoretical Biology* **220**, 169–188.

KOOLHAAS, J. M., KORTE, S. M., DE BOER, S. F., VAN DER VEGT, B. J., VAN REENEN, C. G., HOPSTER, H., DE JONG, I. C., RUIS, M. A. W. & BLOKHUIS, H. J. (1999). Coping styles in animals: current status in behavior and stress-physiology. *Neuroscience and Biobehavioral Reviews* **23**, 925–935.

KWAI, M. (1965). Newly-acquired pre-cultural behavior of the natural troop of Japanese monkeys on Koshima islet. *Primates* **6**, 1–30.

LANDE, R. (2009). Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *Journal of Evolutionary Biology* **22**, 1435–1446.

LEE, A. E. G., OUNSLEY, J. P., COULSON, T., ROWCLIFFE, M. & COWLISHAW, G. (2016). Information use and resource competition: an integrative framework. *Proceedings of the Royal Society London B* **283**, 20152550.

LEFEBVRE, L. (1996). Ecological correlates of social learning: problems and solutions for the comparative method. *Behavioural Processes* **35**, 163–171.

LEVINS, R. (1963). Theory of fitness in a heterogeneous environment. II. Developmental flexibility and niche selection. *American Naturalist* **97**, 75–90.

LIVELY, C. M. (1986). Canalization versus developmental conversion in a spatially variable environment. *American Naturalist* **128**, 561–572.

LOTEM, A. (2013). Learning to avoid the behavioral gambit. *Behavioral Ecology* **24**, 13.

LOTEM, A. & HALPERN, J. Y. (2012). Coevolution of learning and data-acquisition mechanisms: a model of cognitive evolution. *Philosophical Transactions of the Royal Society London B* **367**, 2686–2694.

MAHOMETA, M. J. & DOMJAN, M. (2005). Classical conditioning increases reproductive success in Japanese quail, *Coturnix japonica*. *Animal Behaviour* **69**, 983–989.

MANGEL, M. (1990). Dynamic information in uncertain and changing worlds. *Journal of Theoretical Biology* **146**, 317–332.

MARTIN, J. S. & JAEGGI, A. V. (2022). Social animal models for quantifying plasticity, assortment, and selection on interacting phenotypes. *Journal of Evolutionary Biology* **35**, 520–538.

MATHOT, K. J., WRIGHT, J., KEMPENAERS, B. & DINGEMANSE, N. J. (2012). Adaptive strategies for managing uncertainty may explain personality-related differences in behavioural plasticity. *Oikos* **121**, 1009–1020.

MCDONALD, P. G., KAZEM, A. J. N. & WRIGHT, J. (2007). A critical analysis of 'deceptive' or 'false feeding' behaviour in a cooperative bird: disturbance effects, satiated nestlings or deceit? *Behavioural Ecology and Sociobiology* **61**, 1623–1635.

MCNAMARA, J. M. (1985). An optimal sequential policy for controlling a Markov renewal process. *Journal of Applied Probability* **22**, 324–335.

MCNAMARA, J. M. & HOUSTON, A. I. (1980). The application of statistical decision-theory to animal behavior. *Journal of Theoretical Biology* **85**, 673–690.

MCNAMARA, J. M. & HOUSTON, A. I. (1987). Memory and the efficient use of information. *Journal of Theoretical Biology* **125**, 385–395.

MCNAMARA, J. M. & HOUSTON, A. I. (2009). Integrating function and mechanism. *Trends in Ecology and Evolution* **24**, 670–675.

MCNAMARA, J. M. & LEIMAR, O. (2020). *Game Theory in Biology: Concepts and Frontiers*. Oxford University Press, Oxford.

MERY, F. & BURNS, J. G. (2010). Behavioural plasticity: an interaction between evolution and experience. *Evolutionary Ecology* **24**, 572–583.

MERY, F. & KAWECKI, T. J. (2002). Experimental evolution of learning ability in fruit flies. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 14274–14279.

MERY, F. & KAWECKI, T. J. (2004). An operating cost of learning in *Drosophila melanogaster*. *Animal Behaviour* **68**, 589–598.

MOIRON, M., MATHOT, K. J. & DINGEMANSE, N. J. (2018). To eat and not be eaten: diurnal mass gain and foraging strategies in wintering great tits. *Proceedings of the Royal Society London B.* **285**, 20172868.

MOLDOFF, D. E. & WESTNEAT, D. F. (2017). Foraging sparrows exhibit individual differences but not a syndrome when responding to multiple kinds of novelty. *Behavioral Ecology* **28**, 732–743.

MORAN, N. A. (1992). The evolutionary maintenance of alternative phenotypes. *American Naturalist* **139**, 971–998.

MORAND-FERRON, J., SOL, D. & LEFEBVRE, L. (2007). Food-stealing in birds: brain or brawn. *Animal Behaviour* **74**, 1725–1734.

MURREN, C. J., AULD, J. R., CALLAHAN, H., GHALAMBOR, C. K., HANDELSMAN, C. A., HESKEL, M. A., KINGSOLVER, J. G., MACLEAN, H. J., MASEL, J., MAUGHAN, H., PFENNIG, D. W., RELYEA, R. A., SEITER, S., SNELL-ROOD, E., STEINER, U. K., *ET AL.* (2015). Constraints on the evolution of phenotypic plasticity: limits and costs of phenotype and plasticity. *Heredity* **115**, 293–301.

MURREN, C. J., MACLEAN, H. J., DIAMOND, S. E., STEINER, U. K., HESKEL, M. A., HANDELSMAN, C. A., GHALAMBOR, C. K., AULD, J. R., CALLAHAN, H. S., PFENNIG, D. W., RELYEA, R. A., SCHLICHTING, C. D. & KINGSOLVER, J. (2014). Evolutionary change in continuous reaction norms. *American Naturalist* **183**, 453–467.

NELSON, R. J. & KRIEGSFELD, L. J. (2017). *An Introduction to Behavioral Endocrinology*, Fifth Edition. Sinauer Associates, Sunderland.

NICOLAUS, L. K. & NELLIS, D. W. (1987). The first evaluation of the use of conditioned taste aversion to control predation by mongooses upon eggs. *Applied Animal Behaviour Science* **17**, 329–346.

NUSSEY, D. H., POSTMA, E., GIENAPP, P. & VISSER, M. E. (2005). Selection on heritable phenotypic plasticity in a wild bird population. *Science* **310**, 304–306.

NUSSEY, D. H., WILSON, A. J. & BROMMER, J. E. (2007). The evolutionary ecology of individual phenotypic plasticity in wild populations. *Journal of Evolutionary Biology* **20**, 831–844.

OLIVEIRA, R. F. (2009). Social behavior in context: hormonal modulation of behavioral plasticity and social competence. *Integrative and Comparative Biology* **49**, 423–440.

ORD, T. J., STAMPS, J. A. & LOSOS, J. B. (2010). Adaptation and plasticity of animal communication in fluctuating environments. *Evolution* **64**, 3134–3148.

PAPAJ, D. R. (1994). Optimizing learning and its effect on evolutionary change in behavior. In *Behavioral Mechanisms in Evolutionary Ecology* (ed. L. A. REAL), pp. 133–153. University of Chicago Press, Chicago.

PAPAJ, D. R. & PROKOPY, R. J. (1989). Ecological and evolutionary aspects of learning in phytophagous insects. *Annual Review of Entomology* **34**, 315–350.

PIERSMA, T. & DRENT, J. (2003). Phenotypic flexibility and the evolution of organismal design. *Trends in Ecology and Evolution* **18**, 228–233.

PONZI, E., KELLER, L. F., BONNET, T. & MUFF, S. (2018). Heritability, selection, and the response to selection in the presence of phenotypic measurement error: effects, cures, and the role of repeated measurements. *Evolution* **72**, 1992–2004.

RATIKAINEN, I. I. & WRIGHT, J. (2013). Adaptive management of body mass in Siberian jays. *Animal Behaviour* **85**, 427–434.

READER, S. M. & LALAND, K. L. (2003). *Animal Innovation*. Oxford University Press, Oxford.

ROWLEY, I. (1977). Communal activities among white-winged choughs *Corcorax melanorhamphus*. *Ibis* **120**, 1–20.

SCHEINER, S. M. (1993). Genetics and evolution of phenotypic plasticity. *Annual Review of Ecology, Evolution, and Systematics* **24**, 35–68.

SCHEINER, S. M. (2006). Genotype-environment interactions and evolution. In *Evolutionary Genetics: Concepts and Case Studies* (eds C. W. FOX and J. B. WOLF), pp. 326–338. Oxford University Press, New York.

SCHEINER, S. M. (2013). The genetics of phenotypic plasticity XII. Temporal and spatial heterogeneity. *Ecology and Evolution* **3**, 4596–4609.

SCHLICHTING, C. D. & PIGLIUCCI, M. (1998). *Phenotypic Evolution: Reaction Norm Perspective*. Sinauer Associates, Inc, Sunderland.

SCHULTE, P. M., HEALY, T. M. & FANGUE, N. A. (2011). Thermal performance curves, phenotypic plasticity, and the time scales of temperature exposure. *Integrative and Comparative Biology* **51**, 691–702.

SEEBACHER, F., WHITE, C. R. & FRANKLIN, C. E. (2015). Physiological plasticity increases resilience of ectothermic animals to climate change. *Nature Climate Change* **5**, 61–66.

SHETTLEWORTH, S. J. (1984). Learning and behavioural ecology. In *Behavioural Ecology: An Evolutionary Approach*, Second Edition (eds J. R. KREBS and N. B. DAVIES), pp. 170–194. Blackwell Scientific, Oxford.

SHETTLEWORTH, S. J. (2010). *Cognition, Evolution & Behavior*, Second Edition. Oxford University Press, Oxford.

SIH, A. (2013). Understanding variation in behavioural responses to human-induced rapid environmental change: a conceptual overview. *Animal Behaviour* **85**, 1077–1088.

SIH, A. & DEL GIUDICE, M. (2012). Linking behavioural syndromes and cognition: a behavioural ecology perspective. *Philosophical Transactions of the Royal Society London B* **367**, 2762–2772.

SIMMONDS, E. G., COLE, E. F. & SHELDON, B. C. (2019). Cue identification in phenology: a case study of the predictive performance of current statistical tools. *Journal of Animal Ecology* **88**, 1428–1440.

SIMONS, M. T. T. P., SUVERKROPP, B. P., VET, L. E. M. & DEMOED, G. L. (1992). Comparison of learning in related generalist and specialist eucolid parasitoids. *Entomologia Experimentalis et Applicata* **64**, 117–124.

SKINNER, B. F. (1948). 'Superstition' in the pigeon. *Journal of Experimental Psychology* **38**, 168–172.

SNELL-ROOD, E. C. (2013). An overview of the evolutionary causes and consequences of behavioural plasticity. *Animal Behaviour* **85**, 1004–1011.

SNELL-ROOD, E. C. & STECK, M. K. (2019). Behaviour shapes environmental variation and selection on learning and plasticity: review of mechanisms and implications. *Animal Behaviour* **147**, 147–156.

SOL, D., DUNCAN, R. P., BLACKBURN, T. M., CASSEY, P. & LEFEBVRE, L. (2005). Big brains, enhanced cognition, and response of birds to novel environments. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 5460–5465.

STADDON, J. (2016). *Adaptive Behavior and Learning*, Second Edition. Cambridge University Press, Cambridge.

STAMPS, J. A. & FRANKENHUIS, W. E. (2016). Bayesian models of development. *Trends in Ecology and Evolution* **31**, 260–268.

STAMPS, J. A. & GROOTHUIS, T. G. G. (2010). Developmental perspectives on personality: implications for ecological and evolutionary studies of individual differences. *Philosophical Transactions of the Royal Society of London B* **365**, 4029–4041.

STEPHENS, D. W. (1987). On economically tracking a variable environment. *Theoretical Population Biology* **32**, 15–25.

STEPHENS, D. W. (1989). Variance and the value of information. *American Naturalist* **134**, 128–140.

STEPHENS, D. W. (1991). Change, regularity, and value in the evolution of animal learning. *Behavioral Ecology* **2**, 77–89.

STEPHENS, D. W. (1993). Learning and behavioral ecology: incomplete information and environmental predictability. In *Insect Learning: Ecological and Evolutionary Perspectives* (eds D. R. PAPAJ and A. LEWIS), pp. 195–217. Chapman and Hall, New York.

STEPHENS, D. W., BROWN, J. S. & YDENBERG, R. C. (2007). *Foraging: Behavior and Ecology*. University of Chicago Press, Chicago.

STINCHCOMBE, J. R. & KIRKPATRICK, M. (2012). Genetics and evolution of function-valued traits: understanding environmentally responsive phenotypes. *Trends in Ecology and Evolution* **27**, 637–647.

SULLIVAN, K. A. (1988). Age-specific profitability and prey choice. *Animal Behaviour* **36**, 613–615.

TEBBICH, A., STANKEWITZ, S. & TESCHKE, I. (2012). The relationship between foraging, learning abilities and neophobia in two species of Darwin's finches. *Ethology* **118**, 135–146.

TUFTO, J. (2015). Genetic evolution, plasticity, and bet-hedging as adaptive responses to temporally autocorrelated fluctuating selection: a quantitative genetic model. *Evolution* **69**, 2034–2049.

VAN DE POL, M. & VERHULST, S. (2006). Age-dependent traits: a new statistical model to separate within- and between-individual effects. *American Naturalist* **167**, 764–771.

VAN DE POL, M. & WRIGHT, J. (2009). A simple method for distinguishing within- versus between-subjects effects using mixed models. *Animal Behaviour* **77**, 753–758.

VIA, S., GOMULKIEWICZ, R., DE JONG, G., SCHEINER, M., SCHLICHTING, C. D. & VAN TIENDEREN, P. H. (1995). Adaptive phenotypic plasticity – consensus and controversy. *Trends in Ecology and Evolution* **10**, 212–217.

WATSON, R. A. & SZATHMÁRY, E. (2016). How can evolution learn? *Trends in Ecology and Evolution* **31**, 147–157.

WATSON, R. A. & THIES, C. (2019). Are developmental plasticity, niche construction, and extended inheritance necessary for evolution by natural selection? The role of active phenotypes in the minimal criteria for Darwinian individuality. In *Evolutionary Causation: Biological and Philosophical Reflections* (eds T. ULLER and K. N. LALAND), pp. 197–226. MIT Press, London.

WEST-EBERHARD, M. J. (2003). *Developmental Plasticity and Evolution*. Oxford University Press, Oxford.

WEST-EBERHARD, M. J. (2005). Phenotypic accommodation: adaptive innovation due to developmental plasticity. *Journal of Experimental Zoology* **304**, 610–618.

WESTNEAT, D. F., ARAYA-AJOY, Y. G., ALLEGUE, H., CLASS, B., DINGEMANSE, N. J., DOCHTERMANN, N. A., GARAMSZEGI, L. Z., MARTIN, J., NAKAGAWA, S., RÉALE, D. & SCHIELZETH, H. (2020). Collision between biological process and statistical analysis revealed by mean centring. *Journal of Animal Ecology* **89**, 2813–2824.

WESTNEAT, D. F., HATCH, M. I., WETZEL, D. P. & ENSMINGER, A. L. (2011). Individual variation in parental care reaction norms: integration of personality and plasticity. *American Naturalist* **178**, 652–667.

WESTNEAT, D. F., MUTZEL, A., BONNER, S. & WRIGHT, J. (2017). Experimental changes in brood size alter several levels of phenotypic variance in offspring and parent pied flycatchers. *Behavioural Ecology and Sociobiology* **71**, 91.

WESTNEAT, D. F., POTTS, L. J., SASSER, K. L. & SHAFFER, J. D. (2019). Causes and consequences of phenotypic plasticity in complex environments. *Trends in Ecology and Evolution* **34**, 555–568.

WESTNEAT, D. F., WRIGHT, J. & DINGEMANSE, N. J. (2015). The biology hidden inside residual within-individual phenotypic variation. *Biological Reviews* **90**, 729–743.

WETZEL, D. P., MUTZEL, A., WRIGHT, J. & DINGEMANSE, N. J. (2020). Multivariate mixed models reveal novel sources of (co)variation in nestling begging behavior. *Behavioral Ecology* **31**, 960–970.

WOLPERT, L., TICKLE, C., LAWRENCE, P., MEYEROWITZ, E., ROBERTSON, E., SMITH, J. & JESSELL, T. (2011). *Principles of Development*, Fourth Edition. Oxford University Press, Oxford.

WRIGHT, J. (1997). Helping-at-the-nest in Arabian babblers: signalling social status or sensible investment in chicks? *Animal Behaviour* **54**, 1439–1448.

WRIGHT, J. (1998). Helpers-at-the-nest have the same provisioning rule as parents: experimental evidence from playbacks of chick begging. *Behavioural Ecology and Sociobiology* **42**, 423–429.

WRIGHT, J., KARASOV, W. H., KAZEM, A. J. N., GONCALVES, I. B. & McSWAN, E. (2010*a*). Begging and digestive responses to differences in long-term and short-term need in nestling pied flycatchers. *Animal Behaviour* **80**, 517–525.

WRIGHT, J., McDONALD, P. G., TE MARVELDE, L., KAZEM, A. J. N. & BISHOP, C. M. (2010*b*). Helping effort increases with relatedness in bell miners, but 'unrelated' helpers of both sexes still provide substantial care. *Proceedings of the Royal Society London B* **277**, 437–445.

XII. SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1. Learning *versus* sampling model description.
Appendix S2. R code for the simulations.