



**HAL**  
open science

## Habitat selection and the value of information in heterogenous landscapes

Kenneth Schmidt, François Massol

► **To cite this version:**

Kenneth Schmidt, François Massol. Habitat selection and the value of information in heterogenous landscapes. *Oikos*, 2019. hal-01977021

**HAL Id: hal-01977021**

**<https://hal.science/hal-01977021v1>**

Submitted on 18 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Habitat selection and the value of information in heterogenous landscapes**Kenneth A. Schmidt<sup>1</sup> and François Massol<sup>2</sup><sup>1</sup>Dept of Biological Sciences, MS 3131, Texas Tech Univ., Lubbock, TX 79409, USA<sup>2</sup>Univ. de Lille, CNRS, UMR 8198 - Evo-Eco-Paleo, SPICI group, Lille, France**Corresponding author:** Kenneth A. Schmidt, Dept of Biological Sciences, MS 3131, Texas Tech Univ., Lubbock, TX 79409, USA. E-mail: Kenneth.schmidt@ttu.edu**Decision date:** 30-Sep-2018

---

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1111/oik.05722].

**Abstract**

Despite the wide usage of the term information in evolutionary ecology, there is no general treatise between fitness (i.e. density-dependent population growth) and selection of the environment *sensu lato*. Here we 1) initiate the building of a quantitative framework with which to examine the relationship between information use in spatially heterogeneous landscapes and density-dependent population growth, and 2) illustrate its utility by applying the framework to an existing model of breeding habitat selection. We begin by linking information, as a process of narrowing choice, to population growth/fitness. Second, we define a measure of a population's penalty of ignorance based on the Kullback–Leibler index that combines the contributions of resource selection (i.e. biased use of breeding sites) and density-dependent depletion. Third, we quantify the extent to which environmental heterogeneity (i.e. mean and variance within a landscape) constrains sustainable population growth of unbiased agents. We call this the heterogeneity-based fitness deficit, and combine this with population simulations to quantify the independent contribution of information-use strategies to the total population growth rate. We further capitalize on this example to highlight the interactive effects of information between ecological scales when fear affects individual fitness through phenotypic plasticity. Informed breeding habitat selection moderates the demographic cost of fear commensurate with density-dependent information use. Thus, future work should attempt to differentiate between phenotypic plasticity (i.e. acute fear) and demographic responses (i.e. chronic changes in population size). We conclude with a broader discussion of information in alternative contexts, and explore some evolutionary considerations for information use. We note how competition among individuals may constrain the information state among individuals, and the implications of this constraint under environmental change.

**Keywords:** information, population ecology, spatial heterogeneity

**Introduction** - To determine or modify their own or their offspring's phenotype in a variable world, organisms modulate and exploit statistical variation in the environment for fitness advantage (Lewontin and Levin 2007). To do so requires no more than an organism that perceives the statistical variation in environmental conditions and uses adaptive rules to exploit such variation. In other words, organisms are informed.

Evolutionary ecology embraces at least three non-mutually exclusive contexts in which information is explicitly referenced: (1) phenotype bet-hedging, (2) phenotypic plasticity, and (3) the selection of ecological and social environments, *sensu lato*. Context three often has significant repercussions beyond individual fitness, including density-dependent population growth. As Lewontin and Levins (2007) observed, "Organisms juxtapose bits and pieces of the world and so determine what is relevant". Inherently, this process is guided by information in both ecological and evolutionary time. At the landscape scale, Holt (2009) coined this process 'stitching the niche', in reference to the behavioral flexibility (including movement) to avoid 'being in the wrong place at the wrong time' (Fryxell 1997, Schmidt 2004, Jacob et al. 2015). By such means, populations may persist in environments even where the mean growth rate, averaged indiscriminately across spatial units, would predict otherwise (Pulliam and Danielson 1991, Bascombe et al. 2002, Schmidt 2004, Armsworth and Roughgarden 2005).

Despite the potentially critical role of information use on the persistence and performance of populations, the field of population ecology generally lacks a quantitative framework for estimating how informed individuals are, or need to be, on average, to persist in populations and across landscapes; or the alternative view of how poor of an environment (e.g., mean and variance of the landscape) can a population persist in given a specified set of strategies for obtaining information.

To begin to address these gaps, we build a quantitative framework centered on the relationship between information, spatial heterogeneity, and population growth. We begin by constructing the quantitative relationship between information (or its complement, entropy,  $H$ ) and how a population is distributed among various states (Eigen and Winkler 1981). In our examples, these are mutually exclusive breeding sites. As a scalar,  $H$  does not identify the specific distribution of individuals among sites, but it does narrow the ecological consequences: higher information states have both higher mean fitness and higher maximum fitness potential.

Second, we demonstrate how the Kullback-Leibler index ( $S_{K-L}$ ; Akaike 1973) can be used to quantify a population's expected growth rate as a consequence of how individuals are distributed within a landscape with respect to spatial heterogeneity (Kussell and Leibler 2005). Cumulative  $S_{K-L}$  too is a scalar; however, the summed deviances of site-specific  $S_{K-L}$  weighted by site quality provides a robust

estimate of the contribution of information to population growth while also accounting for negative density-dependence inherent in many ecological phenomena. Third, we define the heterogeneity-based fitness deficit of a landscape based on the distribution of expected fitness of unbiased agents within the landscape. This deficit quantifies the minimum information (i.e., deviation from random) necessary for population viability (i.e.,  $\lambda > 1$ ), thereby linking the statistical properties of landscapes to population persistence. Fourth, we apply the above concepts to a previously published agent-based model to quantify the contribution of non-random site use to population growth under density-dependence, and estimate the maximum capacity of a given information-use strategy to exceed the heterogeneity-based fitness deficit. Last, we explore an example of an interaction between information contexts (i.e., choice and plasticity) using the paradigm of the ecology of fear: Information that biases the choice of breeding location has implications for how females adjust their reproductive effort (Julliard 2000). We use this example to contrast 'acute' fear versus 'chronic' fear at the individual and population level, respectively, within an explicit information context. We demonstrate how chronic fear is influenced by predation risk, and the combination of limited information and limited heterogeneity.

**Part I: Concepts: Information and fitness** - Following the logic of the quantitative model presented in Donaldson-Matasci et al. (2010), consider a simplified behavioral interaction between a hypothetical brood parasitic bird and its putative avian hosts. We assume two of the host species will accept the parasite's egg and two species will reject the egg, and we use annual reproductive success, as a proxy for fitness, vis-à-vis different information states of the brood parasite (i.e., host quality or recognition of rejecter species). The issue of the brood parasite maximizing long-term reproductive success (i.e., number of fledged offspring) becomes one of choosing which host species to parasitize. We assume the uninformed brood parasite lays its four eggs evenly across the four host species (i.e., bet-hedging). Two hosts accept the parasite's egg, which kills the host's own offspring, and two hosts reject the egg (we ignore other sources of mortality to keep the analogy simple). Next, consider that prior observation of hosts' behaviors allows the brood parasite to correctly identify one host species as a rejecter. How much does its expected reproductive success increase with this *information*? If it now divides 4 eggs equally among the remaining three host species (1.33 eggs per host species, on average), its young will be rejected from one host and accepted by two hosts, thus producing 2.67 offspring. The ratio of the change in the parasite's expected reproductive success is  $2.67/2 = 1.33$ , which is identical to the increase in the certainty of host quality:  $(1/3)/(1/2) = 1.33$ .

We could further modify this scenario by introducing variation in quality among hosts that influences the survivorship of the brood parasite, without changing the qualitative conclusions, as we

show below. Hence, information may both reduce the number of options that individuals choose among, and quantitatively alter individuals' selectivity among the remaining options. In addition, acquiring information is unlikely to be cost-free as we have assumed in this example. Costs may include the energy needed for maintaining the sensory apparatus, and movement costs. Other costs are incurred through density-dependent scramble competition (Day and Kokko 2015), which we address briefly in the Discussion at both ecological and evolutionary timescales.

For the remainder, we will focus on the latter, building on the relationship between a population's equilibrium information state and density-dependent population growth. We will also briefly consider the case of information cascades. For example, as information increases the survival of young, as in the above example, we expect females to adjust their trade-off between current and future reproductive success (i.e., phenotypic plasticity). We thus foresee a larger synthesis of information across contexts: Choice, bet-hedging (Starrfelt and Kokko 2012, Crean and Marshall 2009) and phenotypic plasticity (Stephens 1989, Pike, et al. 2016).

**A general model** of habitat selection - The habitat selection behavior of an individual can be summarized by the set of probabilities,  $\{P\}$  (i.e., meta-phenotype), of selecting a particular state from among the set of environmental states,  $\{E\}$ . Borrowing from information theory (Shannon and Weaver 1949), Shannon's entropy ( $H = -\sum [p_i \log_2(p_i)]$ ) quantifies uncertainty in terms of the minimum number of binary decisions necessary to narrow a set of options down to a single choice. We use Shannon's framework to quantify the relationship between information and fitness in general.

Consider an individual selecting a breeding site (e.g., territory or ovipositing location). Its strategy is defined by the set of probabilities ( $p_i$ 's) associated with choosing the corresponding  $i^{th}$  state. Let  $\lambda_i$  be the corresponding lifetime expected fitness for an individual in state  $i$ . Assuming one choice of breeding site per individual per generation, long-term mean fitness of a strategy is calculated as the geometric mean:  $\lambda_{mean} = \prod_i \lambda_i^{p_i}$ . In a two-state system, an individual may utilize a single state (e.g.,  $p_1 = 1, p_2 = 0, H = 0, \lambda_{mean} = \lambda_1$ ), utilize both states equally ( $p_1 = 0.5, p_2 = 0.5, H = 1, \lambda_{mean} = (\lambda_1 \lambda_2)^{1/2}$ ), or utilize the two states unequally ( $p_1 \neq p_2, 0 < H < 1, \lambda_{mean} = \lambda_1^{p_1} \lambda_2^{p_2}$ ).  $H$ , therefore, accounts for both the rejection of some states and the biased use of the remaining states. As indicated earlier,  $H$  is a scalar representation of the lack of information. There are multiple ways a population can be distributed among states for a given value of  $H$  (except at the extremes); however, maximum long-term fitness,  $\lambda_{mean}$ , is constrained by information (Fig. 1).

**Density-dependence and the population benefit of information** – Individuals' collective biased site occupancy affects the population growth rate. However, individuals generate negative population density-dependence through non-random resource use and depletion, including the preemption of space. In addition, the presence and performance of conspecifics (e.g., Betts et al. 2008) may generate social information (Danchin et al. 2004, Szymkowiak 2013) that produces positive density-dependence at low population density, known as an information-mediated Allee effect (Schmidt et al. 2015A). A population's equilibrium information state is, therefore, determined through the opposing actions of biased resource use and depletion. Heterospecifics may also generate social information (Seppänen et al. 2007), but we will not consider this further here.

To quantify a population's equilibrium information state, we use the Kullback-Leibler index ( $S_{K-L}$ ; Akaike 1973) given by:

$$S_{K-L} = \sum_i p_i \log \left( \frac{p_i}{q_i} \right)$$

[1]

This index quantifies the summed degree of non-random site use by a population (Kussell and Leibler 2005), regardless of the source, by measuring the divergence between two distributions (Kussell and Leibler 2005). In our case, these are the realized distribution  $\{q_i\}$  of site occupancy by informed agents within a heterogeneous landscape (i.e., the fraction  $q_i$  of the population occupying site  $i$ ) and the null distribution  $\{p_i\}$  that assumes individuals select sites at random.

The expression given for  $S_{K-L}$  lends itself to an additive decomposition into a sum of site-specific divergence terms, which correspond to the penalty of ignorance components contributed by each site. For instance, if a landscape consists of  $K$  separate sites equally likely to be selected at random, the contribution of site  $i$  to the overall  $S_{K-L}$  is given as  $-\log(q_i K)/K$  where  $q_i$  stands for the probability that an informed organism chooses site  $i$ . The sum of site-specific divergences then yields the overall population penalty of ignorance.

Consider an example of information used in breeding habitat selection that has two sources: information may be (1) based on an individual's prior experience or innate preferences and (2) socially acquired (e.g., conspecific attraction; Danchin et al. 2004). Schmidt et al. (2015A) used a phenomenological model of an information-mediated Allee effect to generate the expected proportions (based on the Wallenius-Hypergeometric distribution) of good ( $\gamma$ ) and poor ( $\pi$ ) sites occupied, and the corresponding per capita growth rate when the population density equals  $N$  (with  $N$  between 0 and carrying capacity,  $K$ ). Assuming the total numbers of good and poor sites are equal to  $G$  and  $P$ ,

Accepted Article  
respectively (with  $G + P = K$ ), and the respective numbers of occupied good and poor sites are given by  $\gamma$   
 $+ \pi = N$ , the corresponding value of  $S_{K-L}(N)$  can be computed as:

$$S_{K-L}(N) = \frac{G}{K} \log\left(\frac{GN}{\gamma K}\right) + \frac{P}{K} \log\left(\frac{PN}{\pi K}\right)$$

In this model, per capita growth rate and  $S_{K-L}$  peak at low population size (Fig. 2) due to the effect of increasing social information, after which the population accumulates an increasing penalty of ignorance due to the depletion of sites, which directly translates into a decline in per capita population growth (Kussell and Leibler 2005). Individuals are informed, but good sites are limiting as density increases. Both contribute to the overall dynamics of population growth.

We have assumed in these examples that density-dependence occurs only through the preemption of sites and the associated contest competition. This may well approximate some taxa, especially territorial vertebrates and certain sessile organisms (Rodenhouse et al. 1997). Other taxa, such as those ovipositing in ponds, can experience more intense density-dependence through scramble competition for limiting resources. In this case, conspecific density within ponds can become the primary source of variation in quality among ponds. Density-dependence may also operate through additional mechanisms external to site selection. These examples do not limit the conclusions we can draw, but rather call for such details to be incorporated into future models.

Below, we use distributions with quasi-continuous site quality and examine both the population  $S_{K-L}$  as well as averages of site-specific divergences within a population over time. Using the latter, we can explicitly quantify the relationship between site occupancy and site quality given different levels (or sources) of information to estimate site bias under density-dependent site use, and its contribution to the population growth rate.

**Information and the heterogeneity-based fitness deficit**– Spatial heterogeneity in local fitness potential has consequences for individuals, regardless of whether they make informed choices. The distribution of expected fitness across habitat patches or landscapes can, therefore, constrain population growth. Here we discuss two related metrics to estimate the effect of landscape heterogeneity on (1) geometric mean population growth and (2) a measure of the minimal information, specifically site occupancy bias, needed to achieve  $\lambda > 1$ . We then demonstrate their use in the following section. This section is motivated by prior results demonstrating that greater variance in site quality requires less habitat selection bias (i.e., information) to achieve  $\lambda > 1$ , and enhances population persistence (Schmidt 2017). However, these findings were shown under the limiting conditions of constant mean site quality.



Here we build on these analyses by developing an estimate of long-term population growth rate based on the unbiased use of landscapes that vary in both the mean and variance in site quality. From this, we quantify the heterogeneity-based fitness deficit, which places limits on the necessary amount of bias in habitat selection to achieve  $\lambda \geq 1$ .

Spatial heterogeneity in site quality can be appreciated in an unbiased way at the landscape level through its distribution (i.e., mean, variance). By contrast, organisms sample site quality through geometric averages as local fitness (dis)advantages accrued from one generation to the next. These considerations can be simply formalized as follows. Defining  $r_i$  as the population growth rate in environmental state  $i$  through  $\lambda_i = e^{r_i}$ , writing  $r$  and  $\lambda$  for the corresponding random variables across the landscape, and noting  $E(X)$  and  $\Pi(X)$  as the arithmetic and geometric averages of quantity  $X$  across the landscape, a simple Taylor series approximation stopping at the second moment of  $\lambda$  yields:

$$\Pi(\lambda) = e^{E(r)} = e^{E(\log \lambda)} \approx e^{\log E(\lambda) - \frac{1}{2} \frac{\text{Var}(\lambda)}{E(\lambda)^2}} = E(\lambda) e^{-\frac{1}{2} \frac{\text{Var}(\lambda)}{E(\lambda)^2}} \approx \frac{E(\lambda)}{\sqrt{1 + \frac{\text{Var}(\lambda)}{E(\lambda)^2}}} \quad [3]$$

The same approximation can be obtained by assuming that  $r$  is normally distributed. In each case, the approximation holds for a semelparous life-history, and when variation in  $\lambda$  is relatively small, i.e., with relatively small variance in fitness across the landscape.

Now consider a landscape that is composed of source sites (sites with net expected fitness,  $\lambda_S > 1$ ) in proportion  $p$  and sink sites (net expected fitness  $\lambda_K < 1$ ) in proportion  $1-p$ . From equation [3], we can estimate the geometric mean  $\lambda_{landsc}$  (where the subscript differentiates  $\lambda_{landsc}$  based on the unbiased occupancy of the landscape versus  $\lambda$  achieved through biased occupancy) based on  $\lambda_S$ ,  $\lambda_K$ , and  $p$ :

$$\lambda_{\text{landsc}} = \frac{p\lambda_S + (1-p)\lambda_K}{\sqrt{1 + \frac{p(1-p)(\lambda_S - \lambda_K)^2}{[p\lambda_S + (1-p)\lambda_K]^2}}} \approx [p\lambda_S + (1-p)\lambda_K] \left[ 1 - \frac{1}{2} \frac{p(1-p)(\lambda_S - \lambda_K)^2}{[p\lambda_S + (1-p)\lambda_K]^2} \right]$$

[4]

Using equation [4], we can set  $\lambda_{\text{landsc}} = 1$  to solve for  $p^*$ , i.e. the minimum proportion of individuals utilizing source sites to achieve  $\lambda = 1$ . This leads to an approximate  $p^*$  value (for  $(\lambda_S - \lambda_K) / 2\sqrt{\lambda_S\lambda_K}$  sufficiently small) given by:

$$p^* \approx \frac{\sqrt{\lambda_K(2 - \lambda_K)} - \lambda_K}{\lambda_S - \lambda_K} + \frac{1}{2} \left( 1 - \frac{\lambda_K}{\sqrt{\lambda_K(2 - \lambda_K)}} \right)$$

[5]

For a given  $\lambda_K$ ,  $p^*$  declines akin to the reciprocal:  $1/(\lambda_S - \lambda_K)$ ; see Fig. 3. The approximation is valid for small  $(\lambda_S - \lambda_K) / 2\sqrt{\lambda_S\lambda_K}$ , i.e. for low differences in local fitness when compared to their geometric average. This approximation can easily hold for non-zero values of  $\lambda_K$  and moderate values of  $\lambda_S$ .

In summary, (1) increasing the variance in site quality, all else being equal, reduces the geometric mean population growth rate of the *unbiased* landscape, while (2) increasing heterogeneity in site quality,  $\lambda_S - \lambda_K$ , requires less biased site use by informed agents to achieve  $\lambda = 1$ . The latter is independent of the proximal mechanisms individuals might use to differentiate between sites, which should also improve under increasing values of  $\lambda_S - \lambda_K$ . Variance and heterogeneity do not quantify the spread in site quality in the same manner. Nonetheless, we note the apparent contrast in their qualitative effects when individuals make informed decisions. Finally, it may also be the case that  $p^*$  and  $\lambda_{\text{landsc}}$  are affected by higher moments of the distributions of local  $\lambda$  values. In the next section we examine the relationships between  $\lambda_{\text{landsc}}$ , the heterogeneity of local  $\lambda$  values, and population persistence of informed agents within randomly generated landscapes. In particular, we use the results above to determine how large of a heterogeneity-based fitness deficit can be overcome by different levels of information to achieve long-term persistence.

**Part II: Application to an agent-based model** - Here we use the agent-based model of Schmidt (2017) to illustrate the utility of the concepts presented above. Agents in this model use the win-stay, lose-switch (WLS) rule of informed site fidelity (Switzer 1993) alone or in combination with conspecific prospecting of qualitative breeding success (e.g., Betts et al. 2008, Kelly and Schmidt 2017) at 4, 8 or 12

sites (specified, for example, as WSLs+4). These strategies play out across a lattice landscape in which breeding site quality is variable in both space and time. Specifically, we consider  $N = 200$  breeding sites, in which the quality ( $Q_i$ ) of the  $i^{\text{th}}$  best site is defined as the probability of success of a single nesting attempt.  $Q_i$  is modeled as  $Q_i = \alpha e^{-\beta(i-1)}$  where the parameters ( $\alpha$ ) and ( $\beta$ ) respectively specify the maximum site quality and the rate of decline in site quality as a function of rank,  $i$ . A higher  $\beta$  increases the relative number of poor sites (i.e., a more right-skewed distribution). Expected fitness from breeding in a site of quality  $Q_i$  is given by  $A_s + \sum \left[ 1 - (1 - Q_i)^3 \right] F$ , where  $A_s (= 0.65)$  and  $F (= 0.6216)$  are the contributions to  $\lambda$  from adult survival and the fecundity of a *successful* nest, respectively,  $Q_i$  is the local probability of success of a single nesting attempt (equal to  $Q_i$  in site  $i$ ), and each individual is allowed two re-nesting attempts following failure.  $F$  is assumed constant among sites. We relax the last assumption in Box 1: 'Chronic fear under informed site-selection' to examine the interactive effects of information in a context of both phenotypic plasticity and selection of the environment.

Density-dependence occurs through the mechanism of the site-dependent population regulation of informed agents (Rodenhouse et al. 1997). Early-settling individuals, when informed, occupy the highest-quality sites, on average, in the absence of prior site preemption. However, increasingly later-settling individuals occupy lower-quality sites due to site preemption. This pattern of biased site occupancy when agents are informed produces site-dependent population regulation (Rodenhouse et al. 1997). Mean site quality in the scenarios we consider yields an expected population growth rate of  $\lambda = 0.97$  under random site selection and without temporal stochasticity in site quality. However, higher per capita population growth rate and long-term population persistence can be realized with information use (Schmidt 2017).

#### **Divergence and the contribution of site bias to population size and growth rate - To**

quantify site bias, we calculated site-specific contributions to divergence ( $S_{K-L}$ ) based on 100 years of site selection and population growth. The total number of site $\times$ year occupied over this time period divided by  $N = 200$  (number of sites) and  $T = 100$  (number of years) was used to generate the expected occupancy under random site selection. We used these values paired with the observed occupancy events (i.e., the observed proportion of years a given site was occupied in the same simulation) to calculate the Kullback-Leibler index for the  $i^{\text{th}}$  site,  $S_{K-L(i)}$  using eq. [3].

Based on the agent-based model described earlier,  $S_{K-L(i)}$  shows no relationship to site quality ( $Q_i$ ) under random site selection (slope = 0.0; Fig. 4A), whereas  $S_{K-L(i)}$  does reveal a strong pattern of under- ( $S_{K-L(i)} > 0$ ; gray) and over-utilization ( $S_{K-L(i)} < 0$ ; colored) of low and high quality sites,

respectively, that is accentuated with greater information use. Due to the inverse symmetry about the median site quality ( $Q_{\text{med}}$ ),  $\sum S_{K-L(i)}$  is relatively uninformative. However, under-utilizing poor sites is a fitness gain relative to random selection, as much as is over-utilizing sites of higher than average quality. We can, therefore, sum each component as positive fitness benefits of non-random occupancy. This in turn provides an approximate measure of the contribution of non-random site use on population growth (see demonstration in the Appendix):

$$\lambda_{\text{contrib}} = \sum_{i=\lfloor N/2 \rfloor + 1}^N \theta_i S_{K-L}(i) - \sum_{i=1}^{\lfloor N/2 \rfloor} \theta_i S_{K-L}(i) \quad [6]$$

where  $\lfloor N/2 \rfloor$  is the median patch index, and  $\theta_i = |Q_i - Q_{\text{median}}|$  is the absolute value of the difference between patch quality and median quality among all patches (using  $Q_{\text{mean}}$  or the x-intercept from the regressions in Fig. 4A-C instead of  $Q_{\text{median}}$  minimally affects the result). Total population growth rate  $\lambda_{\text{total}}$  is then given as the sum of the expected population growth of random agents (0.97) and the contribution of non-random site  $\lambda_{\text{contrib}}$ .

The estimates of total population growth closely match the results of prior population simulations (see Schmidt 2017). Under maximum information use,  $\lambda_{\text{total}}$  is estimated at 1.004, which corroborates the results of stochastic simulations in which ~90% of simulations avoid quasi-extinction (i.e. situations in which  $N < 20$ ) beyond 1000 years (Schmidt 2017). In other words,  $\lambda_{\text{total}}$  is sufficiently high that quasi-extinction is very unlikely over a long-term time horizon (Fig. 5). The relationship between  $\lambda_{\text{total}}$  and the number of information sources  $ns$  (nesting outcomes observed) fits nicely ( $R^2 = 0.94$ ) to the power law:  $\lambda_{\text{total}} = 0.9816 \times (0.00801)^{ns}$  (Fig. 5).

**Overcoming the heterogeneity-based fitness deficit** - The results in the last section estimate the contribution of information use to population growth under the operation of density-dependence, in which population growth does not exceed 1.0 at equilibrium. Hence it underestimates the capacity for information to increase population growth. In addition, population growth will be constrained by variation in landscape heterogeneity, and we wish to understand the limits of those constraints under different levels of information gathering. We examined both of these questions using the agent based-model of Schmidt (2017). To produce variation in the distribution of site quality, we simulated 500 random landscapes in which  $\alpha$  and  $\beta$  were each drawn from a truncated (non-negative) normal distribution:  $\alpha \sim N(0.65, 0.05)$  and  $\beta \sim N(0.121, 0.001)$ . For each landscape, we assessed population persistence time as

the number of years until quasi-extinction ( $N < 20$ ), up to a maximum of 1000 years. We fit persistence time against  $\lambda_{landsc}$  using a three-parameter (a,b,c) sigmoid function with persistence time =

$a / (1 + be^{c\lambda_{landsc}})$  to obtain  $\lambda_{landsc}$  at half the maximum persistence time (i.e., half-saturation point). We consider this as a proxy for the minimum  $\lambda$  necessary to achieve long-term population viability,  $\lambda_{viable}$ .

$\lambda_{viable}$  for agents using the WSLS strategy is 1.005 (Fig. 6B). Below this value, most simulated populations become extinct in  $< 200$  years, whereas above this value, most populations achieve long-term-persistence (Fig 6A, B). Measured against the baseline for long-term persistence of  $\lambda = 1.0$  through the unbiased exploitation of the landscape, we conclude that WSLS alone makes at best a small contribution to long-term population persistence under the specified spatiotemporal stochasticity (a conclusion also reached in Schmidt 2017). In contrast,  $\lambda_{viable} = 0.942$  for agents using the WSLS + 8 prospects strategy. Hence, measured against the same baseline as used above ( $\lambda = 1.0$ ), the WSLS+8 prospecting strategy increases population growth rate by an estimated 0.058.

**Discussion** - The framework of the ecology of information (Schmidt et al. 2010) is based on two universal properties of life: Organisms possess (1) a range of developmental, physiological, and behavioral strategies to perceive the statistical variation in their external environment, and (2) a set of adaptive rules to exploit or modulate the statistical variation in their environment for fitness advantage (Lewontin and Levins. 2007). Exploiting environmental variation, however, often produces density-dependent consequences to population growth, which behavioral, physiological, and developmental contexts often do not consider.

Moreover, in a population context, it is the information state (H) of the population that is of interest. However, because a population is composed of many individuals, there are many configurations (i.e., individuals occupying unique breeding sites in our example) that yield the same information state, but not the same fitness (i.e., population growth rate). Nonetheless, maximum and minimum density-dependent population growth rates are constrained by information. Furthermore, a population's information state is itself density-dependent, which can obscure our ability to infer information use in populations or the benefits thereof. To remedy this, we have utilized the Kullback-Leibler index to quantify the divergence between a population's distribution across breeding sites and random occupancy, and in so doing quantify the contribution of information use strategies to the population growth rate.

Our analyses also quantify a critical influence of spatial heterogeneity on placing limits on *the capacity of information to increase the population growth rate* (Schmidt 2017). This differs from the

potential benefits of spatial heterogeneity, per se (e.g., Bascompte et al. 2002). Rather, the average value of information (Stephens 1989, Pike et al. 2016) increases with greater heterogeneity, and therefore requires a smaller proportion of individuals occupying good sites to achieve  $\lambda > 1$ , all else equal (Fig. 3). This has important implications for how we manage information-scapes beyond the considerations of reducing information-disrupting mechanisms (Vos et al. 2006, Holt 2007, Lüring et al. 2007).

Spatial heterogeneity and temporal variability of the environment may further limit or expand ecological and evolutionary niches within alternative contexts and scales (Orians et al. 1991, Wiens et al. 1993). For instance, foraging individuals and groups use similar information-based strategies to those discussed here for fine and coarse-grain selection of patches within and across multiple scales (see Merkle et al. 2014A, 2015). Likewise, Calcagno et al.'s (2013, 2014) analyses of the marginal value theorem have demonstrated that patchily distributed resources produce two distinct and simultaneous effects on the optimal movement rates of foragers: an average effect and a variance effect. Reducing environmental quality, for instance, produces a variance effect by increasing the coefficient of variation among patches, which in turn increases optimal movement rates between patches. In other words, foragers become cream-skimmers, rejecting poor patches, and reducing time in rich patches that result in higher movement rates. While Calcagno et al.'s (2013, 2014) context differs from breeding habitat selection, an analogous measure of the distribution of patch quality to  $\lambda_{landscape}$  used herein must constrain foragers' optimal patch residence time and thus movements if they are selective on resource patches. Like Calcagno et al. (2014), we suggest that the scaling of variance in site (and patch) quality to mean site quality is a critical property of heterogeneity within habitats, with important consequences for population vital rates.

The ecological benefits of information use may also extend over evolutionary time, although we did not consider it presently. For instance, decreased environmental quality (mean), spatial heterogeneity (variance), and capacity to discern good sites from poor (e.g., poorer signal detection) should intuitively have detrimental consequences for individuals and populations. However, the frequency-dependent advantage for an individual to preempt space, such as a breeding territory, earlier than its competitors can have a significant impact on both population and evolutionary dynamics (Day and Kokko 2015). For example, Schmidt et al. (2015B) showed that such a preemptive advantage results in reduced levels of information gathering commensurate with population size. However, when this preemptive advantage relaxes due to declining demographic rates (e.g., survivorship, brood size), the selective pressure on information gathering increases. As a result, the ESS (i.e., evolutionarily stable strategy) probability of obtaining a good territory increases within a deteriorating habitat (Schmidt et al. 2015B). Hence,

information-gathering may be an important evolutionary mechanism buffering populations from further declines.

Spatial heterogeneity (i.e.,  $\lambda_{landsc}$ ) may also constrain evolutionary buffering and a population's realized niche under information use. Schmidt (2015) analyzed a game-theoretic model in which informed agents used sequential search to locate and occupy breeding sites at or below their ESS maximum predator abundance ( $P^*$ ). Increasing the variance in predator abundance increased the prey's (i) population size,  $N^*$ , (ii) population viability in ecological time, and (iii) expanded the prey's evolutionary niche, such that prey populations persisted with more lethal predators. If these results extend widely to other contexts, environmental heterogeneity coupled with information use may help ameliorate environmental change. Therefore, it is incumbent upon ecologists to recognize the importance of characterizing species information niches, fraught with difficulty as this may be.

## Data deposition

Data available from the Dryad Digital Repository: <<http://dx.doi.org/10.5061/dryad.xxxxx>> (Schmidt and Massol 2018).

**Acknowledgements** - We thank our coauthors on the phenology game (Jacob Johansson, Niclas Jonzén, and Nadiyah Kristensen) for insights on the evolutionary dynamics of information use. We dedicate this manuscript and our continued work on the ecology of information to the memory of Niclas Jonzén: our friend and colleague whose ideas and passion for science still burn bright and continue to inspire us. FM is funded by the CNRS and the ANR-funded ARSENIC project (grant no. 14-CE02-0012).

## References

- Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. - In E. Parzen et al. (eds), Selected Papers of Hirotugu Akaike. Springer Series in Statistics.
- Armsworth, P. R. and Roughgarden, J. E. 2005. The impact of directed versus random movement on population dynamics and biodiversity patterns. - *Am. Nat.* 165: 449 – 465
- Bascompte, J., H. et al. 2002. Patchy populations in stochastic environments: critical number of patches for persistence. - *Am. Nat.*, 159:128-137.
- Betts, M. G. et al. 2008. Social information trumps vegetation structure in breeding-site selection in a migrant songbird. - *Proc. Roy. Soc. Lond. B* 275: 2257-2263.
- Calcagno, V. et al. 2014. The functional response predicts the effect of resource distribution on the optimal movement rate of consumers. - *Ecol. Lett.* 17:1570-1579.
- Calcagno, V. et al. 2013. How optimal foragers should respond to habitat changes: a reanalysis of the Marginal Value Theorem. - *J. Math. Biol.* 69:1-29.
- Cohen D. 1966. Optimizing reproduction in a randomly varying environment. - *J. Theor. Biol.* 12:119–129.
- Crean, A. J. and D. J. Marshall. 2009. Coping with environmental uncertainty: dynamic bet hedging as a maternal effect. - *Phi. Trans. Roy. Soc. Lond B*364:1087-1096.
- Danchin, E., et al. 2004. Public information: from nosy neighbors to cultural evolution. - *Science* 305:487-491.
- Doligez, B., et al. 2002. Public Information and Breeding Habitat Selection in a Wild Bird Population. - *Science* 297:1168-1170.
- Doligez, B., et al. 2003. When to use public information for breeding habitat selection? The role of environmental predictability and density dependence. - *Anim. Beh.* 66:973-988.
- Day, E. and H. Kokko. 2015. Relaxed selection when you least expect it: why declining bird populations might fail to respond to phenological mismatches. - *Oikos* 124:62-68.
- Dillon, K. G. and C. J. Conway. 2017. Nest predation explains variation in avian clutch size. **Behav. Ecol.** 29 301-311.
- Donaldson-Matasci, M. C., et al. 2010. The fitness value of information. - *Oikos* 119:219-230.
- Eigen, M and R. Winkler 1981 *Laws of the game: How the principles on nature govern chance.* Princeton University Press.



- Holt, R. D. 2007. IJEE soapbox: the unraveling of nature's information webs: the next depressing frontier in conservation? *Israeli Journal of Ecology and Evolution* 53:229-236.
- Holt, R. D. 2009. Bringing the Hutchinsonian niche into the 21st century: Ecological and evolutionary perspectives. - *Proc. Nat. Acad. Sci.* 106 (Suppl. 2):19659-19665.
- Jacob, S., et al. 2015. Habitat matching and spatial heterogeneity of phenotypes: implications for metapopulation and metacommunity functioning. - *Evol. Ecol.* 29: 851-871.
- Kelly, J. K. and K. A. Schmidt. 2017. Fledgling calls are a source of social information for conspecific, but not heterospecific, songbird territory selection. - *Ecosphere* 8:e01512.10.1002/ecs2.1512
- Kivelä, S. M., et al. 2014. The past and the present in decision-making: the use of conspecific and heterospecific cues in nest site selection. - *Ecology* 95:3428-3439.
- Kokko, H. 2007. *Modeling for field biologists and other interesting people*. Cambridge University Press, Cambridge, UK.
- Kussell, E. and S. Leibler. 2005. Phenotypic diversity, population growth and information in fluctuating environments. - *Science* 309:2075-2078.
- LaManna, J. A. and T. E. Martin. 2016. Costs of fear: Behavioral and life-history responses to risk and their demographic consequences vary across species. - *Ecol. Lett.* 19:403-413.
- Lewontin, R. and R. Levins. 2007. *Biology Under the Influence: Dialectical Essays on Ecology, Agriculture, and Health*. Monthly Review Press, NYC, NY.
- Lürling, M. and Scheffer, M. 2007. Info-disruption: pollution and the transfer of chemical information between organisms. - *TREE* 22:374-379.
- Martin, T.E. 1995. Avian life history evolution in relation to nest sites, nest predation, and food. *Ecol. Monogr.*, 65, 101–127.
- May, R.M. 1974. *Stability and Complexity in Model Ecosystems*. Princeton University Press. Princeton, NJ.

- Merkle, J.A., et al. 2015A. Bison distribution under conflicting foraging strategies: site fidelity versus energy maximization. **Ecol.** 96:1793–1801.
- Merkle, J.A., et al. 2015B. To follow or not? How animals in fusion-fission societies handle conflicting information during group decision-making. **Ecol. Lett.** 18: 799-806.
- Merkle, J.A., et al. 2014. A memory-based foraging tactic reveals an adaptive mechanism for restricted space use. **Ecol. Lett.** 17:924-931.
- Orians, G. H. and Wittenberger, J. F. 1991. Spatial and temporal scales in habitat selection. – *Am. Nat.* 137: S29 – S49.
- Pike, R. K., et al. 2016. A general expression for the reproductive value of information. - *Behav. Ecol.* 27:1296-1303.
- Pulliam, H. R. and Danielson, B. J. 1992. Sources, sinks, and habitat selection: A landscape perspective on population dynamics. - *Am. Nat.* 137: S50-S66.
- Rodenhouse, N. L., et al. 1997. Site-dependent regulation of population size: a new synthesis. - *Ecology* 78: 2025-2042.
- Schmidt, K. A. 2017. Information thresholds and habitat loss and population persistence in breeding birds. - *Oikos* 126: 651-659.
- Schmidt, K. A. 2014. Site fidelity curbs sequential search and territory choice: a game theoretic approach. - *Funct. Ecol.* 28:1494-1503.
- Schmidt, K. A. 2004. Site fidelity in temporally correlated environments enhances population persistence. - *Ecol. Lett.* 7:176-184.
- Schmidt, K.A. and C. J. Whelan. 2010. Nesting in an uncertain world: information and sampling the future. - *Oikos* 119: 245-253.
- Schmidt, K. A. and Massol, F. 2018. Data from: Habitat selection and the value of information in heterogenous landscapes. – Dryad Digital Repository, <<http://dx.doi.org/10.5061/dryad.xxxxx>>.
- Schmidt, K. A., et al. 2015A. Information mediated Allee effects in breeding habitat selection. - *Am. Nat.* 186: E162 – E171.
- Schmidt, K.A., et al. 2010. The ecology of information: an overview on the ecological significance of making informed decisions. *Oikos* 119:304-316.
- Schmidt, K.A., et al.. 2015B. Consequences of information use in breeding habitat selection on the evolution of settlement time. - *Oikos* 124:69-80.

- Seppänen, J.-T., et al. 2007. Social information use is a process across time, space, and ecology, reaching heterospecifics. - *Ecology* 88:1622-1633.
- Shannon, C. E. and W. Weaver. 1949. The mathematical theory of communication. The University of Illinois Press, Urbana, IL.
- Starrfelt, J. K. and H. Kokko. 2012. Bet-hedging – a triple tradeoff between means, variances and correlations. - *Biol. Rev.* 87:742-755.
- Stephens, D. W. 1989. - Variance and the value of information. *Am. Nat.* 134: 128-140.
- Vos, M., et al. 2006. Infochemicals structure marine, terrestrial, and freshwater food webs: implications for ecological informatics. *Ecological Informatics* 1:23-32.
- Wiens, J. A. et al. 1993. Ecological mechanism and landscape ecology. – *Oikos* 66: 388 – 380.
- Zanette, L. Y. et al. 2011. Perceived predation risk reduces the number of offspring songbirds produce per year. - *Science* 334:1398-1401.

**Box 1. Acute fear versus chronic fear under informed site-selection** - Consider the following life-history principle: When spatial variation results in some breeding sites being less prone to nest failure than others, informed individuals should alter the number and/or quality of offspring among locations that differ in the risk of failure (Julliard 2000). Therefore, informed females, relative to ignorant females, should (1) overproduce offspring when breeding in sites that have higher than average nest survival, and (2) under-produce offspring (i.e., avoid life-history trade-offs for doomed young) in sites that have lower than average nest survival (e.g., Martin 1995, Dillon and Conway 2017).

To investigate the population consequences of this plasticity, we simulated informed agents (using WLSLS+8 strategies) that adjusted (fearful) or did not adjust (fearless) fecundity as a function of the site-specific probability of nest success,  $Q_i$ . We varied the range of spatial heterogeneity in  $Q_i$  within 1000 random landscapes by introducing randomly drawn parameters that govern the distribution of site quality ( $\alpha$  and  $\beta$ ; see above). Under fear, the fecundity of a successful nest at site  $i$  was adjusted as:

$$F_i = F_{max} - cE[A_i] \quad [B1]$$

where  $F_{max}$  (=0.6216) is maximum fecundity,  $E[A_i]$  is the expected number of nest attempts (i.e., =  $1/Q_i$ ) necessary to produce one success, and  $c$  is the rate at which fecundity declines with  $E[A_i]$ . As above, individuals were limited to three breeding attempts to produce one success. Setting  $c = 0.025$ , mean and median adjusted fecundity are reduced by 27% and 21%, respectively, compared to Zanette et al's (2011) estimate of a 40% reduction.

This moderate demographic cost of fear results in a marginal increase in per capita reproductive success (~ 1%) in fearful populations, and a considerable reduction in mean population size (46.2 versus 70.4; Fig. B1). Hence, comparable levels of fecundity may conceal an underlying demographic cost of fear when individuals make informed choices. Instead, fear is manifested through reduced population size, and in that vein reaches a commensurate equilibrium state. We call this chronic fear to differentiate its effects from acute fear quantified by Zanette et al. (2011) and LaManna and Martin (2016). Chronic fear is likely to be more indicative of the combination of limited information and limited heterogeneity (and ensuing density-dependence) rather than be attributed solely to fierce predators (Brown et al. 1999).

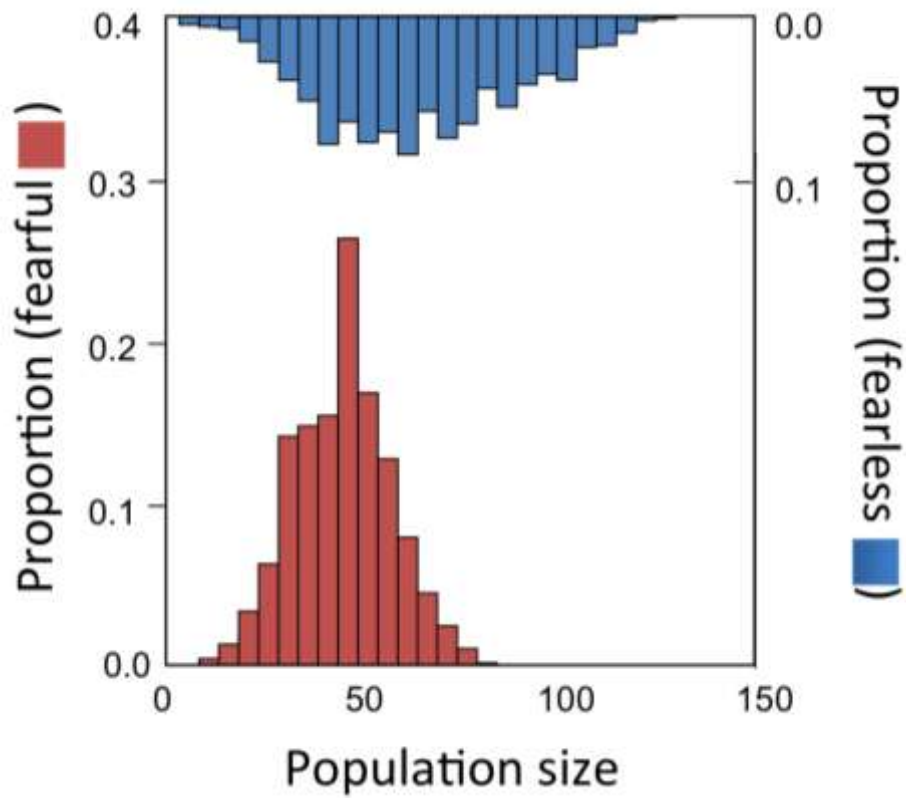


Figure Box 1. Distribution of simulated population sizes of fearful and fearless populations. Fecundity of agents was set to 0.6216 (fearless) determined using equation [B1] (fearful).

## Appendix – Mathematical demonstration of equation (6)

Assume  $q_i$  is patch quality in patch  $i$ ,  $p_i$  is the biased utilization of patch  $i$  and  $1/N$  is its random utilization. The average fitness or patch quality, as experienced by organisms which bias their utilization of patches, is given by:

$$\lambda_{landsc} = \sum_i p_i q_i \quad (\text{A.1})$$

We can partition equation (A.1) as follows:

$$\lambda_{landsc} = \frac{1}{N} \sum_i q_i + \sum_i \left( p_i - \frac{1}{N} \right) q_i \quad (\text{A.2})$$

The first term in equation (A.2) is

$$\lambda_{random} = \frac{1}{N} \sum_i q_i \quad (\text{A.3}),$$

i.e. the average fitness expected without habitat selection. The second term in equation (A.2) will be coined  $\lambda_{contrib}$  and further partitioned as:

$$\lambda_{contrib} = \sum_{i=1}^N \left( p_i - \frac{1}{N} \right) (q_i - q_{med}) + q_{med} \sum_{i=1}^N \left( p_i - \frac{1}{N} \right) \quad (\text{A.4}),$$

where  $q_{med}$  is the median patch quality. In this partition, the second term equals zero because

$$\sum_{i=1}^N \left( p_i - \frac{1}{N} \right) = \sum_{i=1}^N p_i - \frac{1}{N} \sum_{i=1}^N 1 = 0 \quad (\text{A.5}).$$

The first term in equation (A.4) can be further divided based on the median patch quality:

$$\lambda_{contrib} = \sum_{i=1}^{\lfloor N/2 \rfloor} \left( p_i - \frac{1}{N} \right) (q_i - q_{med}) + \sum_{i=\lfloor N/2 \rfloor + 1}^N \left( \frac{1}{N} - p_i \right) (q_{med} - q_i) \quad (\text{A.6}),$$

where  $\lfloor N/2 \rfloor$  is the floored integer closest to the median patch. Now remember that patches are ranked from best to worst. So the first sum in this partition has positive  $q_i - q_{med}$  factors, and if habitat

selection is effective (i.e. good patches are over-utilized), then  $p_i - 1/N$  terms are also positive. By contrast, both terms are expected to be negative or null in the second sum.

Finally, if we define  $Q_i = |q_i - q_{med}|$ , then we obtain:

$$\lambda_c = \sum_{i=1}^{\lfloor N/2 \rfloor} \left( p_i - \frac{1}{N} \right) Q_i + \sum_{i=\lfloor N/2 \rfloor + 1}^N \left( \frac{1}{N} - p_i \right) Q_i \quad (\text{A.7})$$

The link between equations (A.7) and (6) is not obvious, but results from an approximation. Formally, the

$S_{K-L}(i)$  are defined as:

$$S_{K-L}(i) = -\frac{\log(Np_i)}{N} \quad (\text{A.8})$$

When  $N \left( p_i - \frac{1}{N} \right) \ll 1$ , we can approximate equation (A.8) by:

$$\begin{aligned} S_{K-L}(i) &= -\frac{1}{N} \log \left[ 1 + N \left( p_i - \frac{1}{N} \right) \right] \\ &\approx -\frac{1}{N} N \left( p_i - \frac{1}{N} \right) = \frac{1}{N} - p_i \end{aligned} \quad (\text{A.9})$$

Plugging approximation (A.9) into the expression for  $\lambda_{contrib}$ , we get:

$$\lambda_c \approx -\sum_{i=1}^{\lfloor N/2 \rfloor} Q_i S_{K-L}(i) + \sum_{i=\lfloor N/2 \rfloor + 1}^N Q_i S_{K-L}(i) \quad (\text{A.10}),$$

i.e. the result given in equation (6).

### FIGURE LEGENDS

Figure 1. Schematic of an individual's bias ( $p_i$ ; probability of choosing breeding site as a function of site quality) as depicted through a binary decision tree implicit in Shannon's formulation. In the example shown, bias site occupancy is determined by predator signal intensities (depicted by icons in panel A) that correlate negatively with site quality ( $q_i$ ; panel B) and in turn, create a positive correlation between  $p_i$  and  $q_i$ . The relationship between the set of all possible long-term geometric mean fitnesses (calculated from  $p_i$  and  $q_i$ ) and information (i.e.,  $2^H$ ) is shown in panel (C).

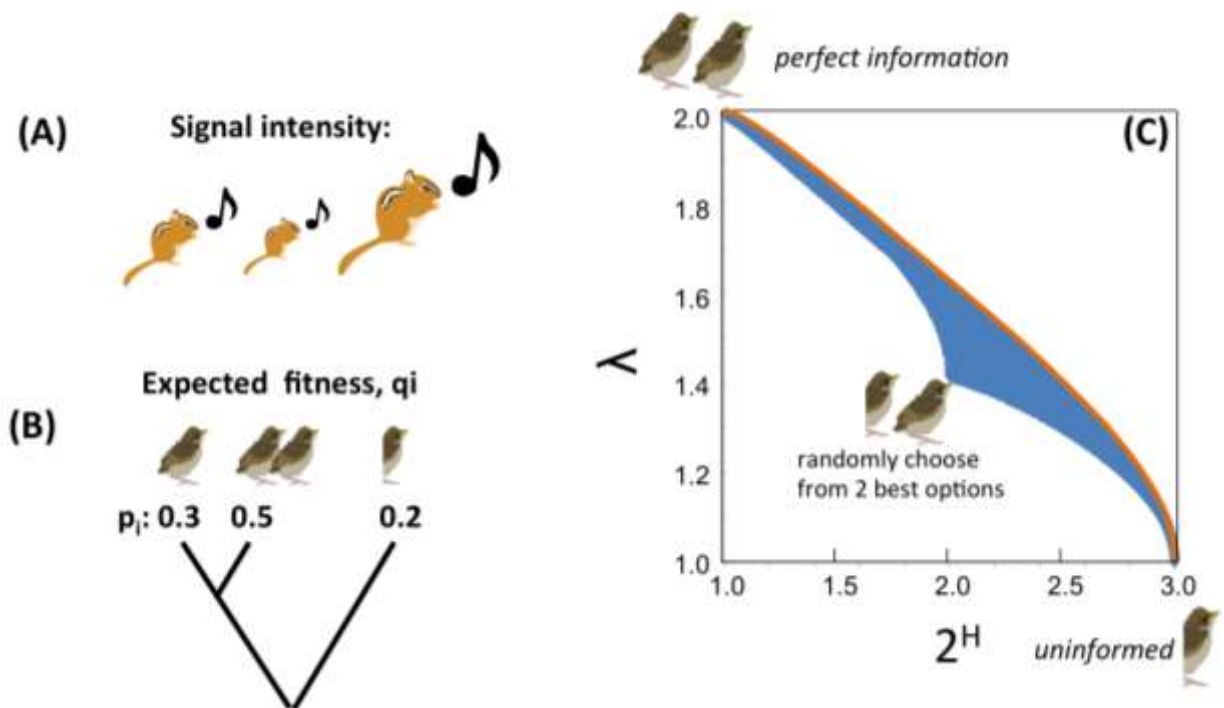




Figure 2. (A) Per-capita population growth rate (pcgr; blue) and Kullback-Leibler divergence ( $S_{K-L}$ ; red) as a function of population size based on the model of an information-mediated Allee effect (see Schmidt et al 2015B).

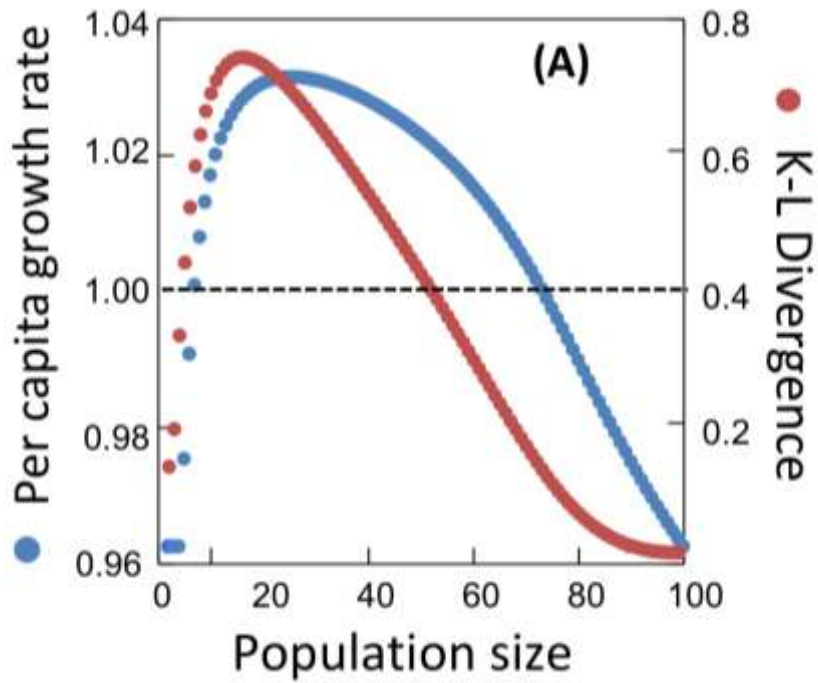


Figure 3. Distribution of  $p^*$  (minimum proportion of individuals utilizing source sites,  $R_S$ , necessary to achieve  $\lambda = 1$ ) is confined to the region in grey and decreases with increasing spatial heterogeneity (i.e.,  $R_S - R_K$ ). For this illustration, the range  $p^*$  was calculated from equation [5] using uniformly distributed values  $1 < R_S < 2$  and  $0.9 < R_K < 0.99$ ). Arrows represent the directional change in  $p^*$  (but not the magnitude) in response to a unilateral increase in  $R_S$  (blue) or decrease in  $R_K$  (red).

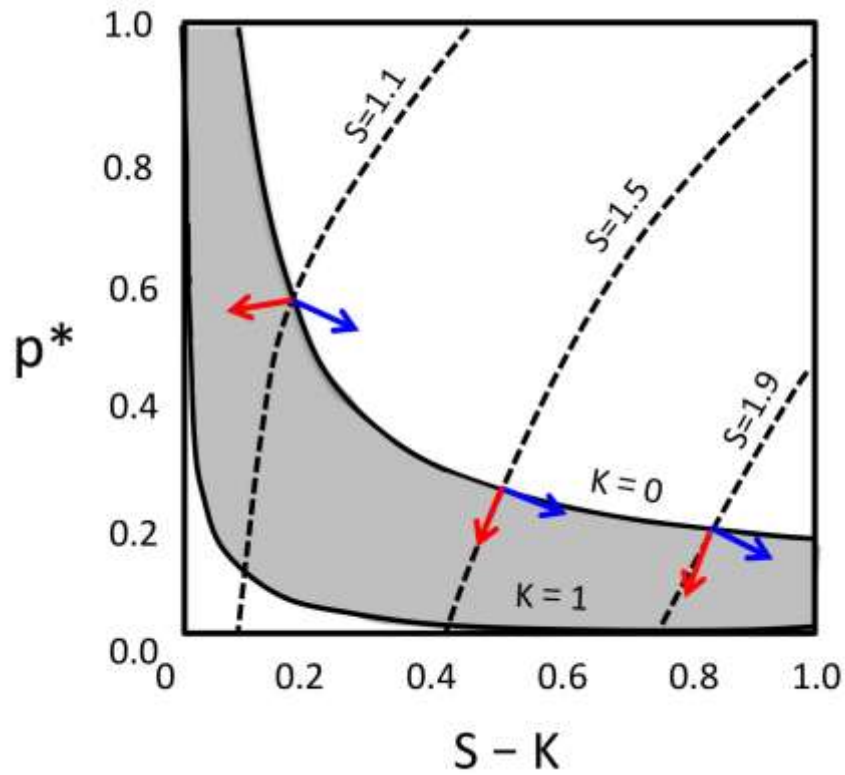


Figure 4. Site-specific Kullback-Leibler divergence versus long-term site quality under three information use scenarios: (A) no information, (B) win-stay:lose-switch rule (WSLS) alone, and (C) WSLS + 8 prospecting events. Slope ( $m$ ) and y-intercept values ( $b$ ) are shown at the top of each panel.

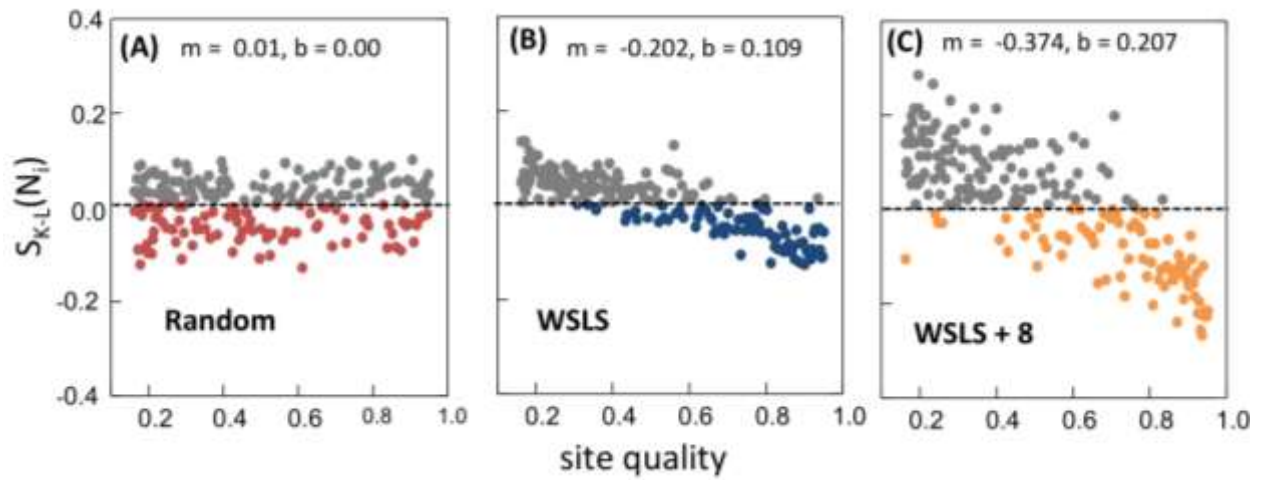


Fig. 5.  $\lambda_{\text{contrib}}$  and total per capita population growth ( $\lambda_{\text{random}} + \lambda_{\text{contrib}}$ ) for the five levels of information use discussed in the text. Values within the panel denote mean population size for the respective information use strategies (see Schmidt 2017).

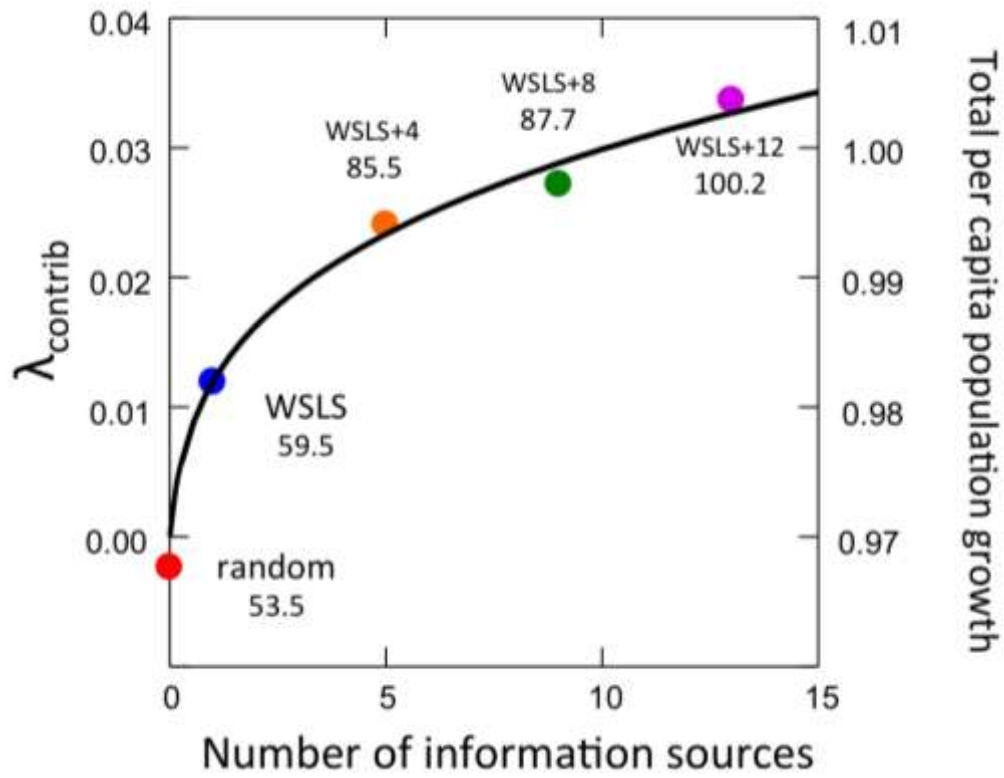


Figure 6. Population persistence time against  $\lambda_{\text{landsc}}$  for two information scenarios: (A) WSLs + 8 prospects and (B) WSLs alone. Each point represents a unique landscape consisting of 200 breeding sites. Variation among landscapes was generated using randomly drawn parameter ( $\alpha, \beta$ ) in the generating function,  $Q_i = e^{(i-1)}$ . Additional model details can be found in Schmidt (2017).

