State-dependent behavioral theory and the evolution of play

Nicholas Grunloh[1] and Marc Mangel[2,3]

Abstract

We show how state-dependent behavioral theory, as implemented by stochastic dynamic programming can enhance our understanding of play, assuming that play is an evolutionary adaptation rather than a non-functional consequence of other processes. We first explain in generic terms the components of a state-dependent behavioral model. We then develop a state-dependent life history model on the assumption that social play is adaptive and thus focus on play's contribution toward the development of skill and how such development affects an individual's future lifetime expected reproductive success (i.e. fitness that accounts for both survival and reproduction). Although our model is not intended to capture a specific situation, a number of commonly observed phenomena emerge from the model, including that most social play is dyadic and that animals select partners that are of similar age or skill. We discuss ways of increasing the fidelity of the model by allowing injury due to play, the role of energetic reserves, handicapping, and polyadic play.

1. Introduction

Burghardt (2005) identifies play according to the following criteria:

(i)   play is behavior that is not immediately necessary for survival (although play may contribute to the development of future fitness);
(ii)   play is a self-motivating behavior; done for its own sake (play is "fun");
(iii) play is often a non-serious version of a similar non-play behavior;
(iv)   play is practiced often, yet loosely stereotyped and experimental in nature; and
(v)   play only occurs in a stress-free environment (the "relaxed field").

These criteria provide a clear framework for the sorts of behaviors that can and cannot be considered play. In addition, they give some sense of just how and when play can occur for the purpose of guiding a model. If we view social play as an evolutionary adaptation, with costs and benefits, then it is appropriate to ask how natural selection will act on play and what the evolutionary outcome of this process is predicted to be. In this paper, we introduce state-dependent behavioral theory (Clark & Mangel, 2000; Houston & McNamara, 1999; Mangel, 2014; Mangel & Clark, 1988; Mangel &

Ludwig, 1992) and show it can be used to model the evolution of play. We specifically choose a simple model that is not specific to any situation but has much in common with many play situations. To do this, we view the benefits of play as the acquisition of skill to be used at some time in the future (e.g. Caro, 1988, 1995; Fagen & Fagen, 2004, 2009). We summarize these benefits in terms of a single quantity, the individual's skill without specifying a mechanism (e.g. maintenance of physical fitness, improved dexterity, or improved social standing). Similarly, the costs associated with play can be loosely grouped into manageable quantities such as the costs associated with not playing (e.g. not maintaining physical fitness) and the costs which occur while playing (e.g. injury or mortality). The observation that play occurs in the presence of these costs suggests that the benefits of play outweigh the costs. Thus, we assume play behavior has evolved because the benefits

[1]Graduate Program in Statistics and Applied Mathematics, University of California, Santa Cruz, CA, USA
[2]Center for Stock Assessment Research, Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA, USA
[3]Department of Biology, University of Bergen, Bergen, Norway

Corresponding author:
Marc Mangel, Center for Stock Assessment Research, Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA 95064, USA.
Email: msmangel@ucsc.edu

of play are greater than the costs, making play behavior an evolutionary adaptation; see Burghardt (2005) for evidence.

If play is adaptive in this way, as opposed to a coincidental non-functional behavior, then the development of skill during play early in life will contribute to expected reproductive success (fitness) later in life and play decisions are predicted to follow a pattern of increasing an organism's fitness through skill (i.e. decisions associated with play should be in some sense be tuned to increases fitness). That is, even though individuals are driven to play because it is "fun" the functional interpretation as to why play has become "fun" is that play at a given period of development increases an organism's fitness at some time in the future (Burghardt, 2005; Caro, 1988; Fagen & Fagen, 2004, 2009).

In the next section, we first describe the components of a state-dependent behavioral model, including characterizing the environment, the states and their dynamics, the definition of fitness in terms of expected lifetime reproductive success, and some comments about implementation. We then describe a specific model in which play develops fitness enhancing skills. We show how to combine backward iteration of the equation of stochastic dynamic programming (SDP) and forward Monte Carlo simulation to make behavioral predictions. We conclude with a discussion of these results and ways that the model could be extended, with the intention of stimulating a deeper connection between empirical studies and modeling.

2. Methods

We begin with a description of the components of a state-dependent behavioral model. We then model a situation in which play allows the development of skills that contribute to subsequent lifetime reproductive success. Our model is not intended to be "realistic", in the sense that we are not modeling a particular play situation by a particular animal. However, we expect that the model has much in common with many different organisms in many different situations, and thus may inspire other colleagues to bring specific details to this modeling framework.

2.1 The components of a state-dependent behavioral model

2.1.1 The environment. In general, we must consider acquisition of food and the risk of acquiring that food (some older readers may recall the "crisis of the common currency" in behavioral ecology in the late 1970s and early to mid-1980s; see Mangel (2014)). Clearly, such risks are minimized for juveniles whose parents provide food and shelter, which is the situation that we consider here.

2.1.2 States and their dynamics. Organisms have physiological states (e.g. body mass, length, fat reserves, skill associated with neuro-muscular development) that change with time, in response to the environment and behavior. For simplicity, the only state that we consider is skill developed through play.

2.1.3 The fitness increment and lifetime fitness. Biology is well suited for economic thinking because there is a natural pay-off from behavior: the representation of genes in future generations. Often, a proxy is used such as accumulated lifetime reproductive success or the number of grand offspring (Clark & Mangel, 2000; Mangel, 2006, 2014; Mangel & Clark, 1988). In the case of a juvenile organism, the situation is simpler because by definition a juvenile is not accumulating reproductive success. Thus, at the end of the juvenile period, we can assess future expected reproductive success as a function of state. The question is then: how do we associate play behavior during the juvenile period with future fitness? The equations of SDP allow us to formalize this question mathematically and derive many predictions, both quantitative and qualitative, that both can be tested empirically and provide insight into the biological world.

2.1.4 Implementation. As will be seen, deriving the equation of SDP forces one to think deeply about the biology of play. Often, some kinds of preliminary analysis can be conducted on the model. However, numerical solution of the SDP equation can provide exceptional insight: both qualitative patterns and detailed numerical predictions. Indeed, as noted by Mangel and Clark (1988) and Clark and Mangel (2000), very often the intuition from a numerical model can be so powerful that one no longer needs the model to understand the phenomenon. And that, of course, is what we are aiming for: understanding the natural world.

2.2 Juvenile play develops fitness-enhancing skills

In order to simplify the dynamics of social play in the model, we consider a focal individual separately from all of the other potential play partners in the environment (see the Discussion section for future directions changing this assumption). Individuals can have skill levels ranging from a minimum skill, S_L, to a maximum skill, S_U (see Table 1 for all symbols and their interpretation).

At time t, the focal individual has level of skill $S(t)$, with particular value noted by i. Similarly, potential

Table 1. Variables, parameters, and their interpretation.

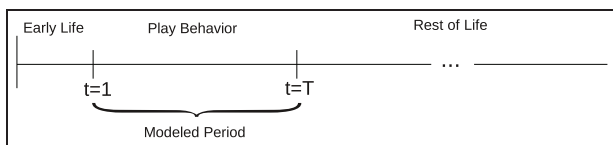| Symbol	Interpretation	Value
$S(t)$	Skill of focal individual at time t	See the text
S_L	Minimum level of skill	0
S_U	Maximum level of skill	50
i	Particular value of skill level	Varies
j	Particular value of the skill of a possible play partner	Varies
α	Per-period loss of skill	0.9
T	Time at which development of skill stops	40
$\phi(i)$	Expected lifetime reproductive success for an individual with $S(T) = i$	Equation (4), Figure 2
$F(i, t)$	Fitness function for an individual with $S(t) = i$	Equation (1), Figure 2
$\Delta S(i, j)$	Gain in skill when an individual whose whose skill is i plays with an individual whose skill is j	Equation (2), Figure 3
S_{max}	Maximum value of increment in skill	20
σ	Characterization of the width of the increment in skill	Varies, Figure 3
τ	Length of a play bout	3
$\lambda_j(t)$	Probability focal individual encounters a play partner with skill level j	Equation (3)
c	Scale parameter for the encounter distribution	0.09
δ_n	Normalization constant for the encounter distribution	0.083
λ_0	Probability of not encountering a play partner in a single period of time	0.05
γ	Shape parameter for $\phi(i)$	Varies
S_0	Values of skill at which $\phi(S_0) = 0.5$	25
$V_{cont}(i, t)$	Fitness value for an individual with skill level i at time t continuing to play	Equation (6)
$D_e^*(i,t)$	Decision of an individual with skill level i at time t to exit the play field or not	Output of equation (5)
$D_p^*(i,j,t)$	Decision of an individual with skill level i at time t to play with an individual of skill level j or not	Output of equation (6)
$R(i, t)$	Range of potential play partners for an individual with skill level i at time t	Emergent, Figure 5

Figure 1. By definition a juvenile is not accumulating reproductive success. However, at the end of the juvenile period, which lasts from $t = 1$ to $t = T$ in our model, future expected reproductive success is assessed according to the level of skill. The equations of SDP allow us to formalize the link between state variables, future fitness, and behavior during the juvenile period.

Figure 2. Three possible choices for terminal fitness of an individual with skill level i, $\phi(i)$. Note the greater the steepness parameter γ the more quickly and dramatically terminal fitness increases for skill greater than S_0 (see equation (4) for details).

play partners have particular skill levels denoted by j. Each time period of the model, the skill of the focal individual decrements by α to capture the idea that skill requires maintenance through repeated practice. At each time period, the focal individual may encounter a play partner, or not; if a partner is encountered, then the focal individual decides to play or not.

We assume that ontogenic development of the skill stops at time T (Figure 1), at which time an individual with skill level i has future lifetime fitness $\phi(i)$. This future fitness is expected reproductive success, that is, the combination of surviving to age and reproduction at that age summed over all ages (see Mangel (2014) for a simple example). We do not model this stage of life explicitly, but rather summarize it through the a function $\phi(i)$, which we show in Figure 2, give mathematically in equation (4), and call the terminal fitness function. Similarly we do not model how skill is
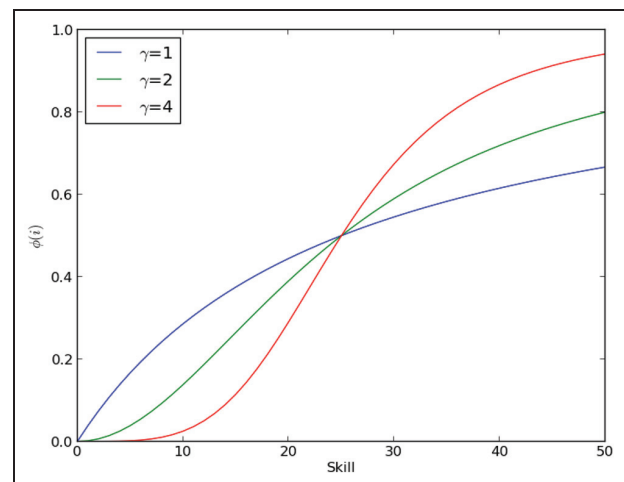
maintained after the ontogenetic period (i.e. for times greater than T), but see the Discussion section.

For previous times we define a fitness function (Clark & Mangel, 2000; Mangel, 2006, 2014; Mangel & Clark, 1988)

$$F(i, t) = \max \ \mathbb{E}[\phi(S(T))] \tag{1}$$

In this equation $\mathbb{E}$ denotes the expectation over encounters with potential partners and "max" denotes

the maximum over the behavioral decision to play or not. Thus, individuals behave adaptively in that they choose whether or not to play based on maximizing their future fitness, not necessarily their immediate fitness. By considering focal individuals with a range of skill levels at any given time, we are able to see how factors independent of energy reserves and predation or injury from play affect an organism's decision to play (see the Discussion section for alternatives).

2.2.1 Play events.
We assume that all play partners are willing and available to enter play events with the focal individual, contingent on the focal individual's decision whether or not to play with them. When a play event occurs between the focal individual, of skill i, and a play partner, of skill j, the focal individual receives an increment to its skill denoted by $\Delta S(i, j)$.

Since skill associated with play events is not necessarily acquired instantaneously, the skill increment of a particular play event is awarded to the focal individual a number of time periods, τ, after the play event starts. Since individuals incur a per-period decrement to their skill, α, every period of the model, and it takes τ time periods to gain skill from a play event, it follows that the total decrement to skill of a single play event is $\alpha\tau$ and the net change in skill when a player of level i plays with a partner of level j is $\Delta S(i, j) - \alpha\tau$. In cases where play events collide with the time horizon of the model, T, we replace $t + \tau$ by T and assume that the focal individual receives the net increment in skill $\Delta S(i, j) - \alpha\tau$. Doing this keeps the relationship between skill increments and skill decrements for truncated play events consistent with all other time periods. However, the effective per period decrement in these truncated play events is higher due to compressing the total cost, $\alpha\tau$, over less than τ time steps.

2.2.2 Skipping play events and exiting the playing field.
The focal individual may skip a play event in a time period because of being unable to find an appropriate play partner, or because the available play partners in the environment do not allow $\Delta S(i, j)$ to be greater than $\alpha\tau$. In this case, the focal individual only incurs the per period cost to skill, α, for a single time period.

Caro's (1988, 1995) results suggest that different types of play occur at differing periods of development and thus a model of play behavior must include the ability of individuals to stop social play altogether. Thus, we include the behavioral option of exiting the play field entirely.

2.2.3 The skill increment more precisely characterized.
We assume that when individuals whose skills are closely matched play, the increment in skill is greater than if the individuals have skills that are widely different (Burghardt, 2006).

Thus, $\Delta S(i, j)$ reaches a maximum, S_{max}, when $i = j$, and as i becomes more different from j, $\Delta S(i, j)$ decreases. For the computation we use the symmetric form

$$\Delta S(i, j) = S_{max} \exp\left\{-\frac{(i - j)^2}{2\sigma^2}\right\} \quad (2)$$

Here σ is a parameter that describes how similar the focal individual must be to the play partner in order to receive a meaningful skill increment from a play event (Figure 3). Here $\Delta S(i, j)$ will always be maximized when the focal individual and the play partner have the same skill (i.e. $i = j$). Note that the symmetry of equation (2) means that $\Delta S(i, j)$ does not really depend on either i or j, but rather the absolute difference between i and j.

As a thought experiment to help understand how focal individuals are motivated by the acquisition of skill through $\Delta S(i, j)$, consider a focal individual that makes play decisions based only on the effects of those behaviors in the short-term. Such a myopic individual only considers whether a play event causes an increase or decrease in skill, regardless of any ill effects these decisions cause in further time periods. For the myopic focal individual the decision to play or not is a comparison between the skill decrement of the play event, $\alpha\tau$, and the skill increment, $\Delta S(i, j)$. If $\Delta S(i, j)$ is greater than $\alpha\tau$ then the myopic individual will always play, regardless of how small the difference, and if $\alpha\tau$ is the greater than $\Delta S(i, j)$, the myopic individual will never play. However, as long as $\tau > 1$ (which we assume it always is), there is an opportunity cost associated with playing with a poorly matched partner. Thus, optimally behaving individuals consider factors that reduce opportunity costs and lead to more selective behavior than in the myopic case.

2.2.4 Play partners.
We characterize the play environment through a probability distribution of potential partners. That is, we let $\lambda_j(t)$ be the probability that a focal individual encounters a potential play partner of
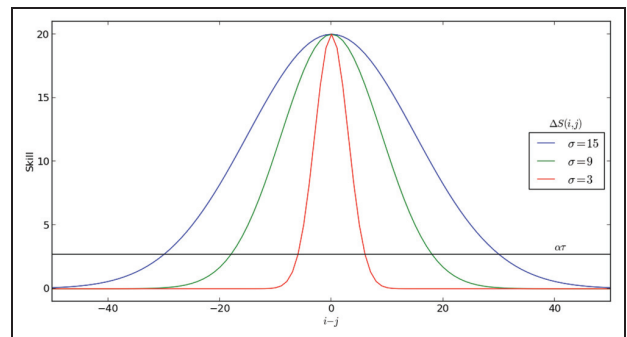

Figure 3. The increment in skill $\Delta S(i, j)$ for player with skill level i playing against a player with skill level j. The horizontal line at $\alpha\tau$ is the decrement in skill during the play interval, so that net gain in skill is $\Delta S(i, j) - \alpha\tau$.

skill j at t. For computations, we use an exponential distribution

$$\lambda_j(t) = \delta_n \exp\{-cj\} \qquad (3)$$

where c is a scale parameter and δ_n is a normalization constant chosen so that $\sum_j \lambda_j(t) \leqslant 1$; $\lambda_0 = (1 - \sum_j \lambda_j(t))$ is the probability that the focal individual cannot find any play partner in a single period of time. The distribution of the potential social play partners in the environment, as an exponential, translates into an environment with initially many low skill individuals. As potential play partners develop their own skills, and leave the population, a decreasing number of high skill individuals are left in the population, unless there is a flow of new individuals into the population (e.g. through births that are asynchronous).

2.2.5 Fitness function and SDP equation.
As described above, we assume that $\phi(i)$ is an increasing function of i, consistent with a higher level of skill at T providing greater future fitness. For computations, we choose a logistic function that has some midpoint skill level after which fitness rises rapidly with skill (Figure 2):

$$\phi(i) = \frac{(i - S_L)^\gamma}{(i - S_L)^\gamma + (S_o - S_L)^\gamma} \qquad (4)$$

Thus, $\phi(i)$ is normalized to a maximum value of 1, S_o is the skill at which half maximal fitness is achieved, and γ characterizes how quickly fitness increases for skills greater than S_o.

In light of the definition of $F(i, t)$ we have the end condition $F(i, T) = \phi(i)$. At each time prior to the time horizon, T, an individual with skill level i may exit the play field, thus obtaining future fitness $\phi(i)$ or may continue to seek partners. We let $V_{cont}(i, t)$ denote the future expected fitness of an individual with skill level i at time t who continues to seek play partners. Then

$$F(i, t) = \max[\phi(i), V_{cont}(i, t)] \qquad (5)$$

The fitness value of continuing to seek play partners depends upon the skill level of the partner encountered at time t, so that

$$
\begin{aligned}
V_{cont}(i, t) = & \left(1 - \sum_j \lambda_j(t)\right) F(i - \alpha, t + 1) \\
& + \sum_j \lambda_j(t) \max[F(i + \Delta S(i,j) - \alpha\tau, t + \tau), \\
& F(i - \alpha, t + 1)] \qquad (6)
\end{aligned}
$$

For example if the focal individual does not encounter a play partner (the first term on the right-hand side of equation (6)) it is not awarded any skill, but still incurs the per period cost to skill, α. If a play partner of skill level j is encountered in period t, with probability $\lambda_j(t)$, the focal individual must decide between

entering a play event or skipping the play event. If the focal individual decides to play, its skill is incremented by $\Delta S(i, j)$ and decremented by α for every period of the play event. Encountering a play partner and choosing not to play has the same fitness effect as having not encountered a partner in the first place. The solution of equations (4)–(6) leads to two matrices of decisions depending upon i, j, and t. The first, $D_e^*(i, t)$, characterizes whether an individual with skill level i at time t exits the play field or not. The second, $D_p^*(i, j, t)$, characterizes whether an individual, of skill i, chooses to play with a partner, of skill j, at time t of the model.

2.2.6 Monte Carlo implementation of play decisions forward in time.
To predict the behaviors of individuals, we use $D_e^*(i, t)$ and $D_p^*(i, j, t)$ to run a Monte Carlo simulation forward through time (Clark & Mangel, 2000; Mangel, 2014; Mangel & Clark, 1988). In particular, we simulate a number of focal individuals independently making optimal play decisions as predicted by equations (5) and (6). We begin by considering K focal individuals in a much larger play field and assign skill levels that are randomly drawn from a uniform distribution on $[S_L, S_U]$. In each time period of the simulation, each of the focal individuals encounters a potential play partner drawn randomly from the probability distribution of encountering potential play partners of skill j. At each potential play encounter the focal individual either enters a play event, skips a play event, or exits the model according to the decision matrices, at the particular i, j, t conditions of the given play encounter.

The simulation follows the following algorithm for each of the K focal individuals.

(1) $t = 0$.
(2) Randomly assign the kth focal individual an initial skill, $I_k(0)$, between S_L and S_U.
(3) Randomly draw a potential play partner skill level, J, from equation (3).
(4) Look up the appropriate play decision, $D_p^*(I_k(t), J, t)$.
(5.1) If the play decision is "exit"; $I_k(t + 1) = I_k(t)$ and $t \to T$.
(5.2) If the play decision is "play"; $I_k(t + \tau) = I_k(t) + \Delta S(I_k(t), J) - \alpha\tau$ and $t \to t + \tau$.
(5.3) If the play decision is "skip"; $I_k(t + 1) = I_k(t) - \alpha$ and $t \to t + 1$.
(6.1) If $t < T$ go to step (3).
(6.2) If $t \geq T$ then $I_k(T) = I_k(t)$.

3 Results

3.1 From the backward equation

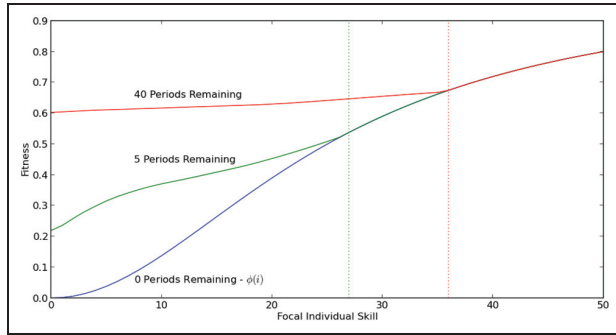Focal individuals choose to play with a range of similarly skilled individuals about the diagonal of $D_p^*(i, j, t)$

Figure 4. The focal individual fitness plotted against skill level. Note that when many time periods remain in the model, fitness is relatively high for all skill levels, due to the prospect of gaining skill in the future. As the number of periods remaining in the model decreases, the fitness of low-skill individuals decreases due to reduced prospect for the future. The dotted vertical lines mark the skill at which the fitness at time *t* converges with the terminal fitness. These dotted lines mark the skill at which the focal individual stops considering play behavior at the given time period of the model. Note that with many time periods of the model remaining only very high-skill individuals exit the model, but as the number of time periods remaining in the model decreases this exit skill decreases.
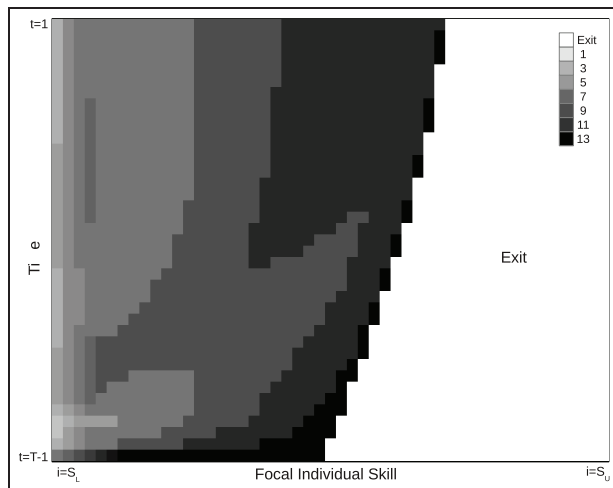

Figure 5. Focal individual play range *R*(*i*, *t*) as a function of both time and focal individual skill level. Dark cells are representative of focal individuals willing to play with play partners of many different skill levels, while light cells are representative of focal individuals with relatively small play ranges. In general, as skill increases, focal individual play range increases. In addition as *t* approaches *T*, in general, play range increases to the myopic condition, at *T* − 1. However, a pocket of lower than expected play ranges violates these general trends. This pocket occurs at relatively high values for *t* and extends across all of the playing skill levels. This pocket is produced by truncating play events as *t* approaches *T*.

where $i = j$. If the cost of play, $\alpha\tau$, is larger than, $\Delta S(i, j)$, a focal individual is predicted not to play with the partner whose skill level is *j*. Thus, $\alpha\tau$ is a major
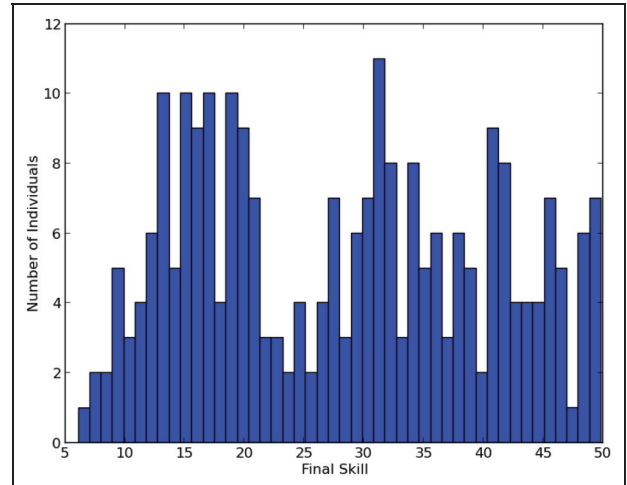

Figure 6. Final skill distribution of *k* = 250 Monte Carlo simulated individuals. Each individual makes optimal decisions, based on $D_p{}^*(i,j,t)$, for 40 time periods. Individuals start the simulation with a random uniformly chosen skill level on the interval [S_L, S_U]. Note the trimodal final skill distribution.

driver in determining the extent to which *i* must be similar to *j* in order for the focal individual to enter a play event. Given a sufficient time horizon, even individuals with low initial skill levels can, through play, develop high final fitness (Figure 4).

We observe patterns in the total range of playable *j*'s based on the focal individuals skill and the time period of the model in which a play event occurs. That is, for each *t* and *i*, there exists a maximum *j* that is beneficial for *i* to play with; we denote this maximum by $\hat{J}_i$. Similarly there is some minimum *j* that is beneficial for *i* to play with, denoted by $\breve{J}_i$. In Figure 5, we show the total range of potential play partners for every combination of *i* and *t*

$$R(i, t) = \hat{J}_i - \breve{J}_i \qquad (7)$$

3.2 Forward Monte Carlo simulation

In Figure 6 we show the final skill distribution of 250 individuals making optimal decisions for 40 periods. This distribution is trimodal. From the left to right, we see: (i) the mode at about skill 15 representing the most common skill for individuals who have not yet exited the model; (ii) the mode around skill 30 representing individuals that begin the simulation below the exit threshold, but play to build enough skill to exit the playing stage; and (iii) the right-most mode around a skill of 40–45, representing individuals that enter the model with skills high enough to automatically exit play behavior.

Another way of visualizing the results of Figure 6 is in a scatter plot (Figure 7), in which we show the relationship between the initial and final levels of skill of
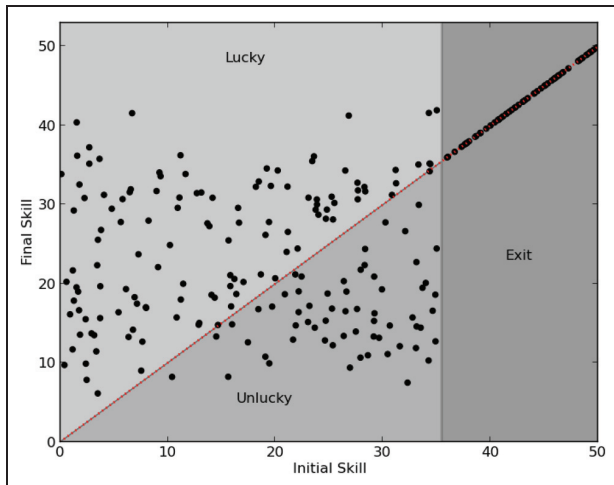
Figure 7. Final skill distribution of $k = 250$ Monte Carlo simulated individuals plotted against the initial skill distribution. The red dotted line indicates the one-to-one relationship between initial and final skill. Individuals on the one-to-one line, in the region labeled "Exit", enter the simulation with high enough skills to immediately exit play behavior.

the simulated individuals. The dotted red 1:1 line shows the final skill level required to maintain the initial skill level. By considering individuals that enter the model with a uniform distribution over the range of possible skills, we ensure that the results will cover all of the possible play strategies in the environment. In such a case, some individuals start with high enough skill to exit the model immediately. These are the individuals with initially high skill, on the one-to-one line in the region labeled "Exit". Individuals with initial skills below the initial exit skill all play to some degree, but the lower the initial skill the more selective the play decisions become. With lower levels of skill, individuals are predicted to be more selective when choosing a play partner. Playing organisms with final skills above the one-to-one line, in the region labeled "Lucky", are individuals that were able to successfully find the play partners that they needed to improve their skill from their initial state. Playing organisms that end up below the one-to-one line, in the region labeled "Unlucky", are individuals seeking play partners, but were not able to find the play partners that they needed to improve their skill. Low-skill individuals are more likely find appropriate play partners and thus they most often end up in the "Lucky" region.

4 Discussion

4.1 General trends

Because we assumed a relaxed play field (*sensu* Burghardt, 2005), we could understand skill-dependent play decisions independent of energy reserves, predation risk, or injury while playing. Although, it is relatively easily to consider play behavior with respect to

these factors, as a starting point it is instructive to understand the basics of play behavior within this simple model first. That is, added considerations may make it hard to see some of the basic forces driving play behavior as seen in this model.

In this model, for simplicity, we have allowed play events between a single focal individual and a single play partner at one time, something which is commonly observed. An emergent property of the model is that individuals preferentially choose play partners with similar skills. This is also observed in natural settings. However, there is no reason that this has to be the case. For example, litters of kittens often play in groups. This may present interesting results considering that the results of this model suggest that playing individuals tend to develop skill as a group. In addition to adding multiple play partners, adding mortality (i.e. survival of a play event is not guaranteed) or injury during play (e.g. leading to a time τ_i without play, and consequent decline in skill) would give insight into a potentially strong cost of play.

When considering the general trends prior to the exception pocket (see below for the exception pocket), we find that low-skill individuals are predicted to be relatively selective in their play decisions, seeking other low-skill individuals. As individuals gain high skill, they become more willing to play with individuals of very dissimilar skill levels. In nature, high-skill individuals may often self-handicap; this could be included in a version of the model in which self-handicapping is another behavioral decision. As shown in Figures 6 and 7, individuals with initially low skill (perhaps the most common natural occurrence) play to increase their skill, and on average they increase their skill level and exit play behavior in the same proportions as other playing individuals. However, one may expect that individuals entering the model with high pre-exit skill levels would have a developmental advantage, and exit the model more quickly and in higher proportions. In general this is not the case, unless playing individuals enter the model virtually at the exit threshold. In general, individuals with initially high pre-exit skill levels quickly fit into very similar skill distributions as individuals with initially low skill. This is due to the scarcity of favorable play partners in the pre-exit upper skill range. On average individuals in a confined social environment will develop their skill as a group. Regardless of an initial skill, the skill development of all individuals in the group converges toward the average skill development of the group.

Individuals with initially very high skill are immediately able to exit. In these cases play behavior is never displayed. This is clearly a hypothetical, and largely unattainable situation for many social species, but these initially exiting individuals could have a meaningful interpretation when one considers behaviors that are not learned via play, or even the evolution of innate behaviors or reflexes.

4.2 The exception pocket

As described above, there is a pocket of time and skill where the general patterns do not hold true. This can be explained by the finite time horizon of the model, and its relation to play events as defined by the model. Recall that for time periods near T, play events cause $t + \tau$ to be greater than T. Due to the construction of the model the skill increments and decrements for play events in these periods are consistent with all other time periods of the model, however the fitness values associated with these skill levels must be truncated at $F(i, T) = \phi(i)$ because by definition fitnesses for time periods beyond T are defined by $\phi(i)$. This has the effect of decreasing $R(i, t)$ for time periods just prior to the final time periods of the model. Skills high enough to exit the model have lower than expected values for $R(i, t)$ several time periods before these individuals exit the model. Individuals several time steps before the end of the model are predicted to be very selective in their choice of play partners because the fitness associated with any skill level in these time periods of the model has been truncated to $F(i, T) = \phi(i)$. Since play is not actually bounded in this way, it is appropriate to run the model with large values of T and consider the general trends prior to this exception pocket.

4.3 Extensions and conclusion

As we have described above, one could include an additional physiological state of energy reserves and potential mortality or injury from play, thus making play more costly. One could also include handicapping, so that a high-skill individual might play with a less-skilled one if there were some benefit either to the focal individual or the play partner (as with mother pandas initiating play with a cub). For simplicity, we treated the play field as exogenous to the behavioral choices of the focal individual. The model could also be extended to those rare species in which a large proportion of their social play involves prolonged polyadic play (e.g. bonobos and some species of macaques). Both of these require a dynamic game between players, in which all individuals decide simultaneously whether or not to play with the encountered potential partner, something beyond the scope of this first paper. Finally, as described above we have not modeled the non-juvenile period subsequent to T. This can be done with the method of sequential coupling (see Clark & Mangel, 2000; Mangel & Clark, 1988). We begin by modeling the period subsequent to T (which we call the reproductive period), in which the end time now corresponds to reproductive senescence, using states in addition to skill (e.g. energy reserves that relate to reproduction) and then couple this model to the one in the current paper so that the first period of the reproductive model corresponds to T of the juvenile model.

In summary, by simplifying the issue to its bare minimum, this state-dependent behavioral model shows an approach that can provide new ways of thinking about the evolutionary biology of play, and much remains to be done.

References

Burghardt, G. M. (2005). *The genesis of animal play: Testing the limits*. Cambridge, MA: MIT Press.

Caro, T. M. (1980). Effects of the mother, object play, and adult experience on predation in cats. *Behavioral and Neural Biology*, *29*, 29–51.

Caro, T. M. (1988). Adaptive significance of play: Are we getting closer? *TREE*, *3*(2), 50–54.

Caro, T. M. (1995). Short-term costs and correlates of play in cheetahs. *Animal Behaviour*, *49*, 333–345.

Clark, C., & Mangel, M. (2000). *Dynamic state variable models in ecology: Methods and Applications*. New York: Oxford University Press.

Fagen, R., & Fagen, J. (2004). Juvenile survival and benefits of play behaviour in brown bears, *Ursus arctic*. *Evolutionary Ecology Research*, *6*, 89–102.

Fagen, R., & Fagen, J. (2009). Play behaviour and multi-year juvenile survival in free-ranging bears, *Ursus arctic*. *Evolutionary Ecology Research*, *11*, 1053–1067.

Houston, A. I., & McNamara, J. M. (1999). *Models of Adaptive Behavior. An Approach Based on State*. Cambridge, UK: Cambridge University Press.

Mangel, M. (2006). *The Theoretical Biologist's Toolbox. Quantitative Methods for Ecology and Evolutionary Biology*. Cambridge, UK: Cambridge University Press.

Mangel, M. (2014) Stochastic Dynamic Programming illuminates the link between environment, physiology, and evolution. *Bulletin of Mathematical Biology*, *77*, 857–877.

Mangel, M., & Clark, C. (1988). *Dynamic modeling in behavioral ecology*. Princeton, NJ: Princeton University Press.

Mangel, M., & Ludwig, D. (1992) Definition and evaluation of behavioral and developmental programs. *Annual Review of Ecology and Systematics, 23*, 507–536.

About the Authors

Nicholas Grunloh received a BS in marine biology in 2012, followed by an MS in statistics and applied mathematics in 2014, from the University of California, Santa Cruz (UCSC) He is currently a Center for Stock Assessment Research (CSTAR) researcher with the National Marine Fisheries Service (NMFS) at the National Oceanic and Atmospheric Administration (NOAA). His research interests include quantitative modeling in behavioral ecology, optimization, and Bayesian hierarchical modeling.

Marc Mangel is distinguished research professor of mathematical biology at UCSC. He conducts research in ecology and evolutionary biology firmly routed in Pasteurs Quadrant in which an applied problem motivates search for basic understanding. He served as the independent scientific expert in the case in the International Court of Justice concerning Japanese special permit whaling in the Southern Ocean.