

University of Groningen

Why do animals have territories?

Hinsch, Martin

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2017

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Hinsch, M. (2017). *Why do animals have territories?*. University of Groningen.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

6 Punish the thief - coevolution of defense and cautiousness stabilizes ownership

Martin Hinsch · Jan Komdeur

Ownership of non-controllable resources usually has to be maintained by costly defense against competitors. Whether defense and thus ownership pays in terms of fitness depends on its effectiveness in preventing theft. We show that if the owners' willingness to defend varies in the population and information about it is available to potential thieves then the ability to react to this information and thus avoid being attacked by the owner is selected for. This can lead to a positive evolutionary feedback between cautiousness in intruders and aggressiveness in owners. This feedback can maintain ownership when the actual direct effectiveness of defense in reducing theft is very low or even absent, effectively turning defense into punishment. We conclude that the deterrence effect of defense in many situations could be stronger than that of prevention and that for many real-world scenarios the purpose of defense of resources might be to punish rather than to drive away intruders.

Submitted to Behavioral Ecology and Sociobiology

6.1 Introduction

Competition for resources such as food or mates is ubiquitous in animals. In order to be able to profit from a resource an individual therefore has to keep competitors away from it (Strassmann & Queller, 2014). Some resources such as small food items can be consumed immediately so that access by conspecifics is easily prevented. For others such as territories or mates, ownership has to be established by means of defense, that is, an aggressive action that reduces a competitor's access to the defended item (Hinsch *et al.*, 2012; Brown, 1964; Maynard Smith & Price, 1973).

Defense is usually costly in terms of time, energy or risk of injury (Schoener, 1983). Whether ownership of a given type of resource is viable in a population therefore depends on whether defense confers a fitness advantage, i.e. whether these costs are lower than the benefit of increased exclusiveness of access to the resource (Brown, 1964). Whether it pays in terms of fitness for the prospective thief to attempt theft depends in turn on the likelihood of being attacked by the owner and the costs of the potentially ensuing fight (Dubois & Giraldeau, 2005). If fighting costs are high and defense (i.e. attack by the owner) is likely enough, therefore, theft can become entirely unprofitable (Hinsch & Komdeur, 2010). This deterrence effect, however, takes place solely in evolutionary time and therefore affects the evolution of defense only indirectly by reducing the average level of attempts of theft and thus the costs of defense.

From studies on the evolution of cooperation we know that deterrence that instead works on the individual time scale can have much more significant effects: If an individual's tendency to punish non-cooperators is known to its competitors before they interact with it they can adjust by being more cooperative towards eager punishers which in turn causes selection for increased punishment. This feedback can be strong enough to lead to the evolution of full "altruistic" cooperation (Johnstone, 2001; dos Santos *et al.*, 2011; Schoenmakers *et al.*, 2014). A similar deterrence effect of aggression has been postulated for dominance hierarchies (Thompson *et al.*, 2014).

Already Stamps (1994) and Stamps & Krishnan (1999) suspected that individual-level deterrence could have similarly dramatic effects on the evolution of resource defense. They even suggested that the main function of aggression towards intruders on territories might be to deter them from intruding again rather than to actively chase them away. It has indeed been shown that if aggression towards intruders is assumed to reduce repeat intrusions, defense by owners is strongly selected for (Switzer *et al.*, 2001). On the other hand we know that it can generally be adaptive for individuals to

avoid returning to the location of a costly encounter (Morrell & Kokko, 2003, 2005). It can therefore be suspected that in a similar way to the co-evolution of punishment and cooperation, sensitivity of intruders for the owner's aggressiveness and willingness to defend could co-evolve to the point where ownership of resources becomes established.

In this study we investigate how the establishment of ownership by co-evolution of defense and theft is affected if potential thieves can obtain and react to information about an owner's willingness to defend. Relative to previous models we make two additional assumptions. First, potential thieves either know or are able to learn by experience the owner's aggressiveness. Second, intruders can evolve sensitivity, i.e. the ability to adjust their tendency to steal from a particular owner based on perceived aggressiveness.

We first introduce a simple mathematical model which allows for a clear demonstration of the general mechanisms involved. In order to verify our results in a more realistic setting we also investigate a more detailed individual-based simulation of a population of territory owners and floaters.

It is important to note that we do not model a particular species and that the floater/territory owner scenario we use for the simulation model is only meant to serve as an example. The evolutionary mechanism we propose is general enough that it could potentially occur in many different real-world situations where animals defend some kind of property against competitors such as defense of feeding territories or mate guarding (van Lieshout & Elgar, 2011; Rousseu *et al.*, 2014; Sherratt & Mesterton-Gibbons, 2015).

6.2 Analytical Argument

We will first use a simple conceptual model to show under which conditions there is selection for defense and sensitivity, respectively.

As our base model we choose a basic sequential move game representing the encounter between an owner and a thief or intruder. The model is structurally similar to others used previously to investigate conflicts between owners and intruders (e.g. Dubois, 2003; Hinsch & Komdeur, 2010; Eswaran & Neary, 2014). For the sake of simplicity, we only model the actual interactions between individuals, leaving out for now details such as resource distribution, search time, frequency of owners and non-owners, etc.

An interaction begins with the intruder deciding whether to attempt theft. If it does not, it has to search for resources elsewhere which we assume results in a frequency dependent (due to competition with other searchers)

Table 6.1: Payoffs in the basic model for owner and intruder, respectively.

owner	intruder	<i>steal</i>	<i>stay away</i>
<i>defend</i>	$V - C_o - S$	$S - C_t$	$P(\hat{t})$
		T	$P(\hat{t})$
<i>concede</i>	$V - T$		V

payoff of $P(\hat{t})$, with \hat{t} as mean population tendency to steal. In this case the owner will gain the full amount V . If the intruder does attempt to steal resources, the owner decides whether to defend or not (following its aggressiveness a). If the owner concedes, the intruder will steal an amount T of resources leading to a reduced payoff of $V - T$ for the owner. If the owner does defend, the intruder only manages to steal a small amount S ($S < T$) resources, however both, intruder and owner have to pay fighting costs C leading to payoffs $S - C_t$ and $V - S - C_o$ respectively. Table 6.1 shows the resulting payoffs for all combinations of actions.

No sensitivity, no variation

Strategies in the basic model - tendency to attempt theft t for the intruder and aggressiveness a for the owner - are modeled as simple probabilities. Given these strategies we can easily spell out the expected payoffs w for a rare mutant intruder (t) and owner (a), respectively, in a homogeneous resident population (\hat{t} and \hat{a}):

$$w_t(t) = (1 - t)P(\hat{t}) + t((1 - \hat{a})T + \hat{a}(S - C_t)) \quad (6.1)$$

$$w_o(a) = (1 - \hat{t})V + \hat{t}((1 - a)(V - T) + a(V - C_o - S)) \quad (6.2)$$

We determine the direction of selection by calculating the gradient of the mutant's fitness around the resident's trait value as $w'_t = dw_t(t)/dt|_{t=\hat{t}}$ and $w'_o = dw_o(a)/da|_{a=\hat{a}}$. An increase in trait value is selected for if the respective derivative is positive (Geritz *et al.*, 1998).

For selection on a we obtain:

$$w'_o = \hat{t}(T - C_o - S) \quad (6.3)$$

It follows that as long as there is any theft ($\hat{t} > 0$) aggressiveness increases if the cost of defense is lower than the benefit of preventing theft:

$$T - S > C_o \quad (6.4)$$

In particular, if defense has no direct prevention effect ($T = S$) then it can never be selected for.

For the intruders we find:

$$w'_t = -P(t) + (1 - a)T + a(S - C_t) \quad (6.5)$$

The propensity to steal correspondingly increases if the net benefit of stealing is greater than the benefit of the outside option (searching for resources elsewhere):

$$(1 - a)T + a(S - C_t) > P(t) \quad (6.6)$$

Since the fitness benefits of aggressiveness are not frequency dependent in this model we will always end up with pure strategies for a , i.e. either 'always attack' ($a = 1$) or 'never attack' ($a = 0$).

For these, we can further simplify inequality 6.6 to either $T > P(t)$ (for $a^* = 0$) or $S - C_t > P(t)$ (for $a^* = 1$). Depending on the choice of function P pure strategies for t are therefore possible if the benefit of the outside option is always lower or always higher than the benefit of stealing irrespective of the frequency of theft. For the purpose of a clear demonstration of the mechanisms involved we will, however, restrict our discussion in the following to the more relevant cases where t has inner equilibria (i.e. $0 < t^* < 1$).

Deterrence selects for aggressiveness

In the next step we investigate the effect of sensitivity for aggressiveness on the evolution of defense. We assume that intruders have a way to know the aggressiveness of an owner they are about to interact with in advance (e.g. by observing conflicts with others or by experience) and are able to modify their behavior accordingly. Therefore, instead of by a fixed probability t to make an attempt at theft the intruders' behaviour is now determined by a sensitivity function $t(a)$ that depends on an owner's aggressiveness a .

We will first investigate how a evolves dependent on the properties of a given (unspecified) function t . Only after that will we take a closer look at the evolution of t itself.

Apart from the change in notation payoffs remain the same as before (see eqn. 6.2):

$$w_o(a) = (1 - t(a))V + t(a)((1 - a)(V - T) + a(V - C_o - S)) \quad (6.7)$$

Which gives us the following selection gradient (with $t' := \frac{dt(a)}{da}$, $w'_o := dw_o(a)/da$):

$$w'_o = t(a)(T - C_o - S) - t'((1 - a)T + a(S + C_o)) \quad (6.8)$$

As before this equation has a straightforward interpretation. The first term on the right hand side is identical (with t replaced by $t(a)$) to the selection gradient in the simple model (equation 6.3) and therefore represents the direct benefits of defense.

In addition, however, as soon as intruders are responsive to the owner's aggressiveness (i.e. as soon as $t' \neq 0$), there is now a second term representing an additional indirect effect of aggressiveness. If intruders are cautious ($t' < 0$) this term becomes positive and even increases with a if defense is costly ($S + C_o > T$). This deterrence effect can thus provide an additional strong benefit to aggressiveness.

For the situation without any direct benefits of defense ($T = S$) that always lead to the disappearance of aggressiveness in the simple model, we obtain now:

$$w'_o = t(a)C_o - t'(T + aC_o) \quad (6.9)$$

If deterrence is strong enough (i.e. intruders are cautious enough) defense can therefore be selected for even if it has no effect on the amount stolen at all.

Variation selects for sensitivity

In the last step we take a look at the evolution of sensitivity, i.e. the ability of intruders to adjust their behavior. Behavioral flexibility can be costly (e.g. Auld *et al.*, 2010; Burns *et al.*, 2011), therefore we do not expect it to evolve in an entirely homogeneous population of owners where sensitivity has no benefit for the intruders at all. In reality, however, aggressiveness will vary between owners due to e.g. mutation, developmental effects such as age or condition or differences in personality (McNamara *et al.*, 2008; Wolf *et al.*, 2008) so that it might be advantageous for intruders to be able to adjust their behavior despite additional costs.

As before we assume that intruders know (by experience or observation) how likely they are going to be attacked when stealing from a specific owner. Let $t(a)$ again denote the probability that an intruder steals dependent on the owner's aggressiveness.

We model variation in a by assuming that with probability p_i intruders will encounter an owner with aggressiveness a_i . For a given population of

owners we obtain the benefit of not intruding as $\hat{P} := P(E\hat{t})$. The cost of sensitivity c_s we assume to be proportional to the variance in t :

$$c_s(t) = C_s \text{VAR}(t(a_i))$$

With this we can formulate our fitness function for the thief:

$$w_t = -c_s(t) + \sum p_i [(1 - t(a_i))\hat{P} + t(a_i)((1 - a_i)T + a_i(S - C_t))] \quad (6.10)$$

Since the trait value t is a vector, calculating the selection gradient as we did it before is not easily possible. Instead, in order to determine the direction of selection, we determine the fitness of a single mutant in a resident population that is assumed to be sufficiently large and homogeneous with respect to the trait under consideration. If the mutant's fitness is positive there is selection for, if it is negative, against that particular mutation. A trait value is evolutionarily stable if no mutant with a higher fitness exists.

If we write the mutant's trait values as $t(a_i) = \hat{t}(a_i) + s_i$ the difference between the mutant's and the residents' fitness Δw_t derives as:

$$\Delta w_t = -\Delta c_s + \sum p_i s_i ((1 - a_i)T + a_i(S - C_t) - \hat{P}).$$

For small mutation step sizes this can be rewritten as:

$$\Delta w_t = \sum p_i s_i ((1 - a_i)T + a_i(S - C_t) - \hat{P} - 2C_s(\hat{t}_i - E\hat{t})).$$

From this we can show (for details see appendix) that there is a single evolutionarily stable strategy in this system that is given by:

$$t_i^* = \frac{(1 - a_i)T + a_i(S - C_t) - \hat{P}}{2C_s} + E\hat{t}.$$

As we see, in the stable state the probability to steal t_i is proportional to the expected fitness payoff from intruding into a territory that is defended with probability a_i . t_i is therefore proportional to the probability not to get attacked by the owner $1 - a_i$ (note that $T > S - C_t$). Any variation in defense a therefore leads to the evolution of an equivalent variation in theft t and thus to the evolution of sensitivity ($t' \neq 0$) and in particular cautiousness ($t' < 0$).

Feedback loop

Since cautiousness in turn increases selection for aggressiveness under certain conditions (see eq. 6.8) a positive feedback loop can occur that drives a coevolution of defense and cautiousness and in this way greatly stabilizes ownership. It is worth noting that cautiousness as well as aggressiveness can be selected for even if defense is never successful in the strict sense, i.e. if it does not reduce the amount of resources an intruder steals ($T = S$).

6.3 Simulation

We tested the results of our mathematical analysis in a more detailed individual-based simulation model. We give a short summary of the model that will be expanded on in more detail below: A population composed of a fixed proportion of territorial owners and non-territorial floaters competes for resources that occur in the territories as well as in an unclaimed area accessible to all floaters. During each time step each floater decides whether to forage either in the unoccupied area (potentially competing with other floaters) or to attempt to intrude into a territory and steal resources. Owners decide whether to start a costly fight in order to attempt to chase away intruders. Fitness of all individuals is determined as sum of resource items foraged minus all fighting costs.

The source code of the simulation program is available from the authors on request.

Evolution

Following common practice we assume haploid parthenogenetic individuals with directly heritable traits (i.e. genotype and phenotype are not distinguished). Each generation of individuals completely replaces the previous generation. An individual's fitness is calculated as overall energy uptake (u , see below) minus costs through fighting and behavioral flexibility (see below). Its expected number of offspring is then determined by the individual's fitness relative to the population mean. On reproduction each evolving trait mutates with a probability of 0.01. Mutation step size is normally distributed with a mean of 0 and a standard deviation of 0.1.

Ecology

At all times the population consists of 1000 individuals. At the start of the simulation and immediately after reproduction half of the population is assigned a territory that they will keep for the rest of their life. The remaining individuals become non-territorial floaters.

Each unit of space is assumed to refill to a level of 1 resource unit at the beginning of each (interaction) time step. The entire habitat has a size of 3000 space units. Territories measure 5 space units while the unoccupied area covers the remaining 500 space units. Unless interrupted (see below) foraging individuals are able to cover 5 space units during one time step. Movements during foraging are assumed to be completely random. If, therefore, several individuals forage in the same area - such as several floaters in the common area or the owner and one or more intruders in a territory - there is a chance that their paths overlap. Given size of the area A and space covered by one individual during foraging f_i and taking overlap into account we can calculate the expected proportion of space visited by at least one individual as

$$F = 1 - \prod 1 - \frac{f_j}{A}. \quad (6.11)$$

Assuming a resource density of 1 unit per space unit, expected resource uptake of an individual i is then

$$u_i = FAf_i / \sum f_j. \quad (6.12)$$

Intrusion and fighting

We assume that floaters have a limited home range that covers 20 territories and does not change during their life time. At the beginning of each time step each floater visits a random territory within its home range and decides whether to attempt to intrude into that territory or whether to forage in the unoccupied area.

In scenarios without sensitivity, floaters intrude with a fixed heritable probability t . In scenarios with flexible behavior individuals adjust their tendency to steal depending on an estimate of the owner's aggressiveness based on their past experience. For each territory floaters keep track of the number of times they have been attacked while intruding on that territory n^a versus the overall number of time steps they spent foraging there n^f . From this the floaters estimate the owner's attack probability as $\tilde{a}_i = n_i^a / n_i^f$. The adjusted tendency to steal $\tilde{t}_i(\tilde{a}_i)$ is then calculated based on traits s and o as:

$$\tilde{t}_i(\tilde{a}_i) = s\tilde{a}_i + o \quad (6.13)$$

A meaningful estimate \bar{a} can only be made after a number of attempts to intrude into the same territory. In scenarios with sensitivity individuals therefore use the sum of t and \tilde{t} , weighted by numbers of intrusion n^f to determine the actual probability p to intrude into a territory:

$$p_i = \frac{n_i^f \tilde{t}_i + t}{n_i^f + 1} \quad (6.14)$$

Behavioral complexity can carry a fitness cost (e.g. for maintenance of the required physiology or increased reaction times, see Auld *et al.*, 2010), therefore floaters pay $C_s |\tilde{t}(0) - \tilde{t}(1)|$ energy units per decision.

After all floaters have made their choice the non-intruding ones move into the common area while the intruders start foraging on the territory they have selected. Owners then decide (according to their trait aggressiveness a) for each intruder on their territory whether to attack or not. Attacks result in costly fights (with costs C_i for intruders and C_o for owners, respectively) that are won by the owner and the intruder with equal probability. If the intruder wins, both, owner and intruder forage with equal efficiency. If the owner wins the intruder is chased away and the area it covers during foraging f (see equations 6.11 and 6.12 above) is reduced to f^L depending on effectiveness of defense e :

$$f_i^L = f_i(1 - e) \quad (6.15)$$

Different values of e could for example be a result of differences in how quickly owners detect intruders and consequently in how much time intruders have for foraging before being detected. Corresponding to the difference between T and S in the mathematical model (see table 6.1), the lower e , the less direct benefits defense has for the territory owner.

The presence of several intruders at once is assumed to have no additional effects beyond fighting costs and potential reduction on foraging efficiency (see section Ecology).

6.4 Results

Unless mentioned otherwise all simulations start out with a peaceful (i.e. non-defending), non-cautious ancestral population and run for 20,000 generations. Results are presented as mean values (and standard error) of 10 replicate runs.

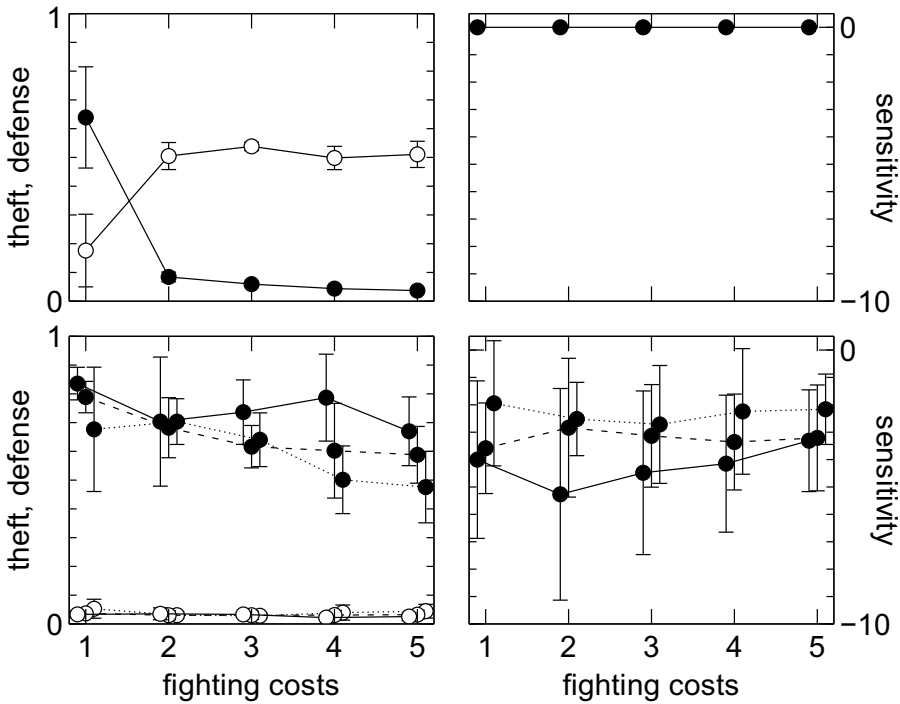


Figure 6.1: *Evolution of ownership with sensitivity set to 0 (top) and evolving (bottom). Defense (left, filled circles), theft (left, lines with open circles at the bottom of the graph) and sensitivity (right) after 20k generations are shown for low (solid line, $c_s = 0$), medium (dashed line, $c_s = 0.05$) and high (dotted line, $c_s = 0.1$) costs of sensitivity. Despite effective defense ($e = 1$) ownership only evolves for low fighting costs if there is no sensitivity. If sensitivity can evolve it does so even for high fighting costs and thus stabilizes ownership.*

Sensitivity

Without sensitivity, ownership only evolves for very low fighting costs (fig. 6.1). For higher costs defense disappears and theft is high.

If intruders can change their behavior based on perceived aggressiveness of an owner then negative sensitivity evolves (fig. 6.1, right). Even a moderate negative slope of tendency to steal versus aggressiveness is sufficient to trigger the evolution of high levels of defense. These in turn lead to very low levels of theft so that ownership becomes established (fig. 6.1).

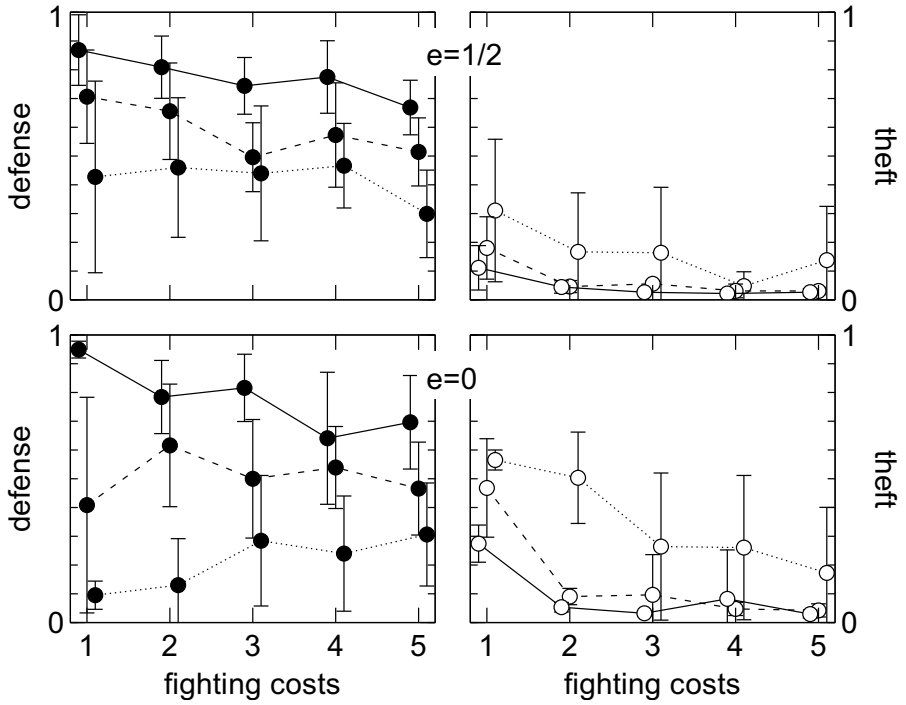


Figure 6.2: Evolution of ownership for moderately effective ($e = 1/2$, top) and completely ineffective defense ($e = 0$, bottom). Defense (left) and theft (right) after 20k generations are shown for low (solid line, $c_s = 0$), medium (dashed line, $c_s = 0.05$) and high (dotted line, $c_s = 0.1$) costs of sensitivity. Even if defense has no direct effect, deterrence can be sufficient to maintain ownership.

Ineffective defense

The mathematical analysis predicted that even if defense has little direct effects it can evolve due to the benefits of deterrence. This is confirmed by our simulation. High levels of defense and consequently low levels of theft can even evolve if owners can not prevent theft at all (fig. 6.2).

Increased variation

Based on the mathematical analysis we expect that an increased variation in attack probability will cause an increase in selection for cautiousness and thus defense. In order to test this we added a small proportion (10%) of tough individuals to the population. As owners these pay only 20% of the fighting costs compared to the rest of the population and consequently can afford to be more aggressive (see Kreps & Wilson, 1982).

As can be seen from figure 6.3 a strong owner disadvantage prevents the evolution of ownership. The presence of a small number of tough individuals, however, increases the variation in aggressiveness experienced by intruders sufficiently to again let cautiousness and defense coevolve to a point where ownership becomes established.

Additional choice

Having floaters choose between one (random) territory and the unoccupied area seems like an artificial restriction given that distances between territories can be small and that individuals have all the information required to make a better decision. We therefore investigated an extension of the model based on the assumption that floaters can choose between intrusion into either of two territories and foraging in the unoccupied area. Each floater is first presented with two random territories (out of its home range). Of those it picks the one with the higher probability to intrude and only then makes the decision whether to intrude or not.

Figure 6.3 shows that adding this extra step to the floater's decision process significantly increases selection for sensitivity.

6.5 Discussion

We have shown that if potential thieves or intruders are capable of adapting their behavior to an owner's aggressiveness in evolutionary time, even small

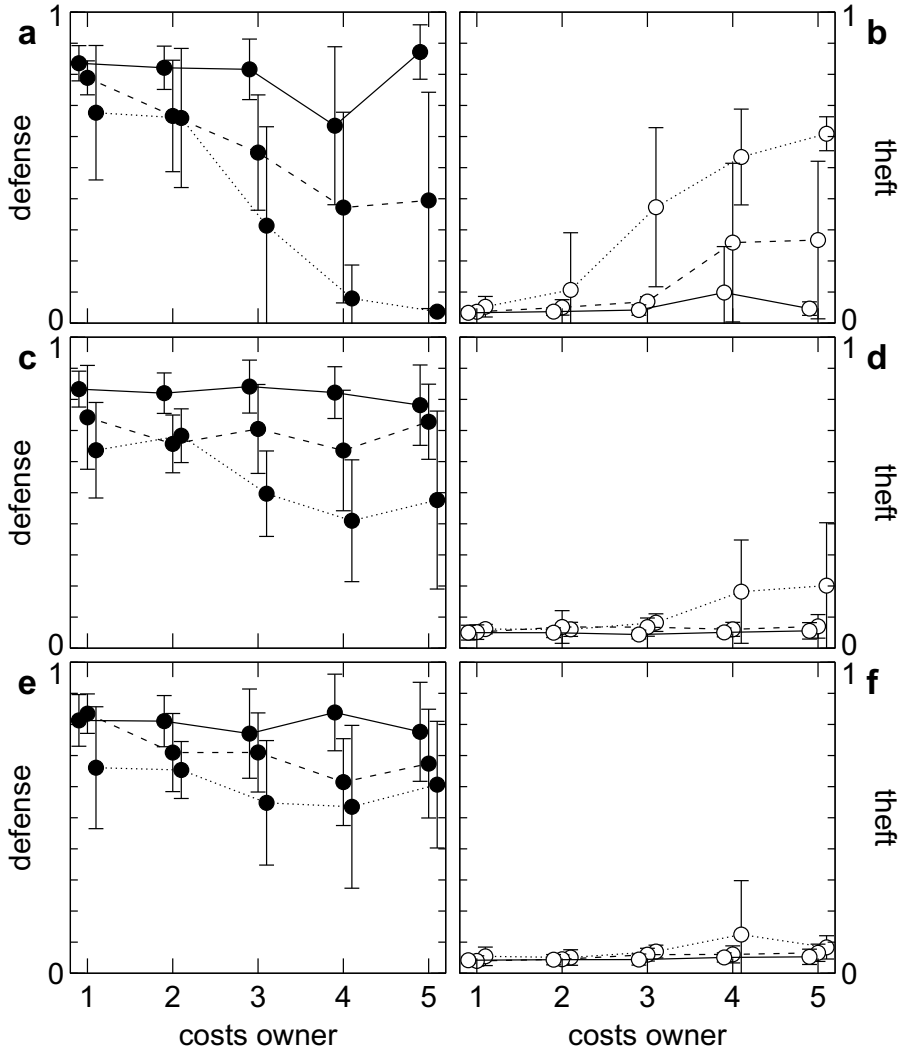


Figure 6.3: *Evolution of ownership for fixed fighting costs for the intruder ($c_i = 1$) and varying costs of defense. Defense (left) and theft (right) after 20k generations are shown for different costs of sensitivity (low: solid line, $c_s = 0$; medium: dashed line, $c_s = 0.05$; high: dotted line, $c_s = 0.1$). Results are presented for the standard scenario (A, B), with a small number of tough owners in the population (C, D) and with intruders choosing the best out of two territories before deciding on intrusion (E, F). A strong owner disadvantage prevents the evolution of ownership for medium and high costs of sensitivity. Higher variation in a as well as additional choice compensate for that. [$e = 1$]*

variations in defense propensity can trigger the evolution of increased cautiousness which again strongly selects for higher aggressiveness and thus leads to the establishment of ownership.

In classical models on the evolution of defense of resources (e.g. Maynard Smith & Parker, 1976; Houston *et al.*, 1985; Morrell & Kokko, 2005; Dubois & Giraldeau, 2005; Gintis, 2007) ownership is viable since intrusion or theft are uneconomical. This is due to the fact that the expected gain from a conflict with the owner (i.e. probability to win times value of resource) is lower than the fighting costs. Although this can be interpreted as a deterrence effect of defense (in particular in models with optional fights, e.g. Dubois & Giraldeau, 2005) it takes place entirely in evolutionary time and has therefore no bearing on the fitness benefits of defense itself. Consequently defense in these models can not be selected for if it has no immediate effect on the intruder's chances of success (Selten, 1978). Similar to Switzer *et al.* (2001) our results show that cautiousness, i.e. flexible reactions to the owner's aggressiveness lead to deterrence in sub-evolutionary time which can increase the benefits of defense up to the point where direct effects are no longer necessary.

Furthermore neither cautiousness nor defense have to be assumed to pre-exist in the population - small random variations are sufficient to trigger a positive feedback between the two traits that leads to the establishment of defense. Defense and ownership can therefore be evolutionarily stable and even emerge in populations that would remain entirely peaceful in the absence of deterrence. This suggests that the feedback between cautiousness and defense can play a much greater role in stabilizing ownership than the immediate effect of defense.

Our results can therefore give a possible explanation for the existence of property in situations where the resource in question is not strictly defendable in the classical sense. Hinsch & Komdeur (2010) for example predicted that in many situations territory owners should profit from poaching on their neighbor's territory to such a degree that territory defense would become untenable in the long term. The existence of deterrence could explain why even in cases where resources would be easily accessible to neighbors only low levels of poaching occur and territoriality is maintained (Carpenter & MacMillen, 1976a; Young & Monfort, 2009; Dantzer *et al.*, 2012).

If defense has no immediate benefit it becomes functionally equivalent to punishment (Raihani *et al.*, 2012). In studies on the evolution of cooperation punishment of cheaters has been proposed as a way that the benefit of unilateral non-cooperation is sufficiently reduced for altruistic behavior to become advantageous in comparison. However since punishment is usually

assumed to be costly it can - in equivalence to ineffective defense - only be selected for if it has some additional positive effects for the punisher (Gardner & West, 2004; Schoenmakers *et al.*, 2014). Similar to our results it has been shown that the availability of knowledge (by reputation or experience) about the individuals' willingness to punish combined with the ability to react to this information can lead to a deterrence effect that is sufficient to compensate for the costs of punishment (Sigmund *et al.*, 2001; dos Santos *et al.*, 2011; Thompson *et al.*, 2014). Together with our results this demonstrates that punishment and defense can be seen as two points on the same continuum.

The occurrence of the described feedback effect in our model rests on a number of conditions concerning physiology and ecology of the modeled species. The generality of our results is determined by how likely it is that these conditions are met in natural populations.

First, individuals have to be able to obtain information about the aggressiveness of their competitors either by personal experience through repeat interaction or through other mechanisms such as direct observation, reputation or signals (but see Hurd, 2006). In any species with either stable social groups or a stable spatial organization this condition will be naturally met (Earley, 2010).

Second, they have to possess the cognitive capabilities to store and use this information. Most vertebrates as well as many invertebrate species are assumed to be capable of at least simple forms of learning (Brembs, 2003). Basic operant conditioning in combination with either spatial memory or individual recognition should be sufficient for the type of information processing assumed in our model (see Gutnisky & Zanutto, 2004; Tanabe & Masuda, 2013).

Third, there has to be sufficient variation in aggressiveness or attack rate to trigger the feedback. While mutation rates in our simulation are relatively high, epigenetic effects as well as environmental stochasticity during development provide additional sources of variation in reality that were not included in the model (see McNamara *et al.*, 2008; Wolf *et al.*, 2008). Our results furthermore suggest that all variations between individuals that lead to variations in attack rate can serve as trigger for the evolution of cautiousness. Besides purely genetic or physiological effects therefore all variation that either produces a phenomenological variation in attack rate (such as detection probability due to e.g. habitat differences) or consistently induces different strategic decisions in different individuals (such as territory quality, individual size, condition or experience) will have the same effect (see Kreps & Wilson, 1982; Przepiorka & Diekmann, 2013). It seems reasonable

to assume that at least some of these sources of variation will be present in most natural populations.

It is also important to note that while at least implicitly our models suggest scenarios with intraspecific competition, there is no intrinsic reason to assume that the same mechanism could not apply to the interaction between individuals of different species such e.g. as interspecific kleptoparasitism (Iyengar, 2008).

In conclusion we think that the conditions for an evolutionary feedback between cautiousness and defense are probably met in many populations in which defense of property occurs. This has ramifications for empirical as well as theoretical research. In empirical studies great effort has been invested to determine the costs and benefits of defense. If a large part of the adaptive value of defense however consists in scaring away competitors from challenging the owner in the first place the measured benefits will necessarily be too low. In most previous models on defense e.g. in the context of mate guarding, territoriality, resource defense or kleptoparasitism only direct benefits of defense have been investigated thereby likely significantly underestimating the range of parameter values for which defense and with it ownership can be evolutionarily stable.

Finally we want to note that a model can only ever be a proof of principle. Whether the mechanism we propose does in fact play a role in a given system can therefore only be determined with the help of empirical research.

Appendix

Using the fitness function from the main text (eq. 6.10)

$$w_t = -c_s(t) + \sum p_i [(1 - t(a_i))\hat{P} + t(a_i)((1 - a_i)T + a_i(S - C_t))],$$

and assuming that a mutant differs in tendency to steal by a small amount s_i we get for the fitness of the mutant:

$$w_{t,m} = -c_s(t+s) + \sum p_i [(1 - (\hat{t}(a_i) + s_i))\hat{P} + (\hat{t}(a_i) + s_i)((1 - a_i)T + a_i(S - C_t))].$$

The fitness difference between mutant and residents then resolves to:

$$\begin{aligned} w_{t,m} - w_t &= -c_s(t+s) + \\ &\quad \sum p_i [(1 - (\hat{t}(a_i) + s_i))\hat{P} + (\hat{t}(a_i) + s_i)((1 - a_i)T + a_i(S - C_t))] + \\ &\quad c_s(t) - \sum p_i [(1 - \hat{t}(a_i))\hat{P} + \hat{t}(a_i)((1 - a_i)T + a_i(S - C_t))]. \end{aligned}$$

This can be simplified as:

$$\begin{aligned}\Delta w_t &= -\Delta c_s + \sum p_i((-s_i\hat{P}) + s_i((1-a_i)T + a_i(S-C_t))) \\ &= -\Delta c_s + \sum p_i s_i((1-a_i)T + a_i(S-C_t) - \hat{P}).\end{aligned}\quad (6.16)$$

As mentioned in the main text we now assume the costs of sensitivity to be proportional to the variance of the values of t (as realized in a given resident population with its associated distribution of aggressiveness a):

$$c_s(t) = C_s \text{VAR}(t)$$

From this we can derive the change in costs:

$$\begin{aligned}\Delta c_s &= C_s(\text{VAR}(\hat{t} + s) - \text{VAR}(\hat{t})) \\ &= C_s(\text{VAR}(\hat{t}) + \sum p_i s_i(2(\hat{t}_i - E\hat{t}) + s_i(1 - p_i)) - \text{VAR}(\hat{t})) \\ &= C_s \sum p_i s_i(2(\hat{t}_i - E\hat{t}) + s_i(1 - p_i)).\end{aligned}$$

Since we assume small mutation steps ($s_i \ll t_i$) the last term in the sum can be neglected, leaving:

$$\Delta c_s = C_s \sum p_i s_i 2(\hat{t}_i - E\hat{t}).$$

Plugging this into Δw (eq. 6.16) we obtain:

$$\Delta w_t = \sum p_i s_i((1-a_i)T + a_i(S-C_t) - \hat{P} - 2C_s(\hat{t}_i - E\hat{t})).$$

For the sake of convenience (and without loss of generality with respect to evolutionary dynamics) we rescale fitness as $w'_t = w_t/2C_s$ and also define the expected (rescaled) payoff of intrusion as

$$I'_i := ((1-a_i)T + a_i(S-C_t))/2C_s.$$

This gives us a simplified expression:

$$\Delta w'_t = \sum p_i s_i(I'_i - \hat{P}' - (\hat{t}_i - E\hat{t})).$$

We can now see that if

$$I'_i - \hat{P}' = \hat{t}_i - E\hat{t} \quad \forall i \quad (6.17)$$

then invasion fitness will always be 0 for any combination of s_i . No mutant can therefore spread in the resident population. On the other hand if we assume a population with a few “mismatched” t_i , so that

$$I'_j - \hat{P}' > \hat{t}_j - E\hat{t}, \quad j \in J$$

and

$$I'_k - \hat{P}' < \hat{t}_k - E\hat{t}, \quad k \in K,$$

then in particular any mutant with $s_j > 0$ and $s_k < 0$ ($\forall j, k$) will have a positive invasion fitness and thus will be able to invade.

We conclude that for a given a the only evolutionarily stable strategy is the vector t that fulfills equation 6.17:

$$\begin{aligned} t_i^* &= I'_i - \hat{P}' + E\hat{t} \\ &= \frac{I_i - \hat{P}}{2C_s} + E\hat{t} \\ &= \frac{(1 - a_i)T + a_i(S - C_t) - \hat{P}}{2C_s} + E\hat{t} \end{aligned}$$

For the sake of completeness we would like to note that since we know that $\sum t_i - Et = 0$ we can derive an explicit solution for t^* that only depends on a :

$$t_i^* = I'_i - EI' + P^{-1}(EI')$$

For the points made in the main text, however, this is not relevant.

