Signalling victory to ensure dominance: a continuous model

Mike Mesterton-Gibbons[1] and Tom N. Sherratt[2]

[1] Mike Mesterton-Gibbons, Department of Mathematics, Florida State University, 1017 Academic Way, Tallahassee, Florida 32306-4510, U.S.A `mesterto@math.fsu.edu`
[2] Department of Biology, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada `sherratt@connect.carleton.ca`

Abstract

A possible rationale for victory displays—which are performed by the winners of contests but not by the losers—is that the displays are attempts to decrease the probability that the loser of a contest will initiate a future contest with the same individual. We explore the logic of this "browbeating" rationale with a game-theoretic model, which extends previous work by incorporating the effects of contest length and the loser's strategic response. The model predicts that if the reproductive advantage of dominance over an opponent is sufficiently high, then, in a population adopting the evolutionarily stable strategy or ESS, neither winners nor losers signal in contests that are sufficiently short; and only winners signal in longer contests, but with an intensity that increases with contest length. These predictions are consistent with the outcomes of recent laboratory studies, especially among crickets, where there is now mounting evidence that eventual winners signal far more frequently than losers after fighting, and that post-conflict displays are more likely to be observed after long contests.

1 Introduction

Bower (2005) defines a victory display as a display performed by the winner of a contest but not by the loser. He offers in essence two possible adaptive explanations of their function: that they are an attempt to advertise victory to other members of a social group that do not pay attention to contests, or cannot otherwise identify the winner, and thus alter their behavior ("function within the network"), or that they are an attempt to decrease the probability that the loser of a contest will initiate a new contest with the same individual ("function within the dyad"). In an earlier paper (Mesterton-Gibbons and Sherratt, 2006), we called the first rationale advertising, and the second one browbeating; and we used game-theoretic models to explore the logic of both rationales. These models showed that both rationales are logically sound; moreover, all other things being equal, the intensity of victory displays will be highest through advertising in groups where the reproductive advantage of dominating an opponent is low, and highest through browbeating in groups where the reproductive advantage of dominance is high.

Here we further consider the browbeating rationale, leaving the case of an advertising rationale for future work. By the browbeating rationale, a victory display is an attempt to decrease the probability that the loser of a contest will initiate a new contest with the same individual. As long as there is a chance that the loser will challenge the winner to another fight in the future, the winner has won a battle for dominance, but not the war. If, on the other hand, the victory ensures that the loser will never challenge, then victory is tantamount to dominance. Thus browbeating is an attempt to ensure that victory equals dominance, and the essence of modelling this phenomenon is to observe a distinction between losing and subordination.

Although we have previously demonstrated that browbeating is a plausible mechanism for victory displays, our earlier model assumed—as opposed to predicted—that a loser does not display, and hence dodged the question of why victory displays should be respected. Moreover, our original model assumed that all contests were of equal length, which leaves open the question as to whether individuals should be more or less likely to signal their dominance after long (close) fights than after short (one-sided) fights.

Accordingly, our purpose here is twofold. First, it is to relax the assumption that the loser does not display. Second, our purpose is also to address the context-dependent nature of the display, which lies outside the scope of our original browbeating model. Specifically, in a recent study investigating fighting behavior in the spring field cricket, *Gryllus veletis*, Bertram et al (2010) found that the intensity of post-conflict signals (aggressive song rate and body jerk rate) were dependent on whether the individual was a winner or loser (with winners signalling more intensely than losers) and on the duration of the contest (with short fights producing less intense signals). Ting et al. (in prep) came to similar conclusions after analysing the outcomes of fights in the fall field cricket, *Gryllus pennsylvanicus*. Likewise, post-conflict displays in the black-capped chickadee, *Poecile atricapillus*—albeit more common among losers than among winners—were more likely to occur after highly aggressive contests (Lippold et al, 2008). Collectively, these recent studies suggest that context dependency might be a general feature of post-conflict displays. Clearly, if mathematical models are to be of value in understanding victory displays, then they should help explain not only the display, but also who displays, and with what intensity. Here we present a simple model that addresses both phenomena.

2 Mathematical Model

The dominance status of a victor relative to the vanquished is determined by a combination of dominance and what we refer to as "non-subordination." A contest outcome is one of dominance if one individual subordinates to the other but the second does not, and of non-subordination if neither defers to the other. To capture the idea that dominance over an opponent contributes more to (long-term) fitness than non-subordination, which contributes more than being dominated, let there be a fitness benefit of 1 for dominating the other individual, of b for non-subordination, and of 0 for being dominated. Thus b, a dimensionless parameter assumed to satisfy

$$0 \ < \ b \ < \ 1 \tag{1}$$

throughout, is an inverse measure of the reproductive advantage of dominance, which is greatest in the limit as $b \to 0$ and least in the limit as $b \to 1$.

We assume that a fight is inevitable; and that its cost can be ignored, by virtue of being the same for both individuals. Neither animal has information about its own or its opponent's strength. So we assume that each animal is equally likely to win, since each has probability $\frac{1}{2}$ of being stronger. However, the length T of the contest is experienced, and the shorter it is, the more likely it is that the winner is the stronger animal.

These assumptions are most nearly satisfied when a fight is provoked between two new neighbors with no prior knowledge of one another's strength and hence no established dominance relation. Of course, they are idealizations. Yet not only do they yield a tractable analytical model, but also they enable us to avoid confounding the evolution of post-conflict displays with the evolution of basic aggression thresholds per se.

Let strategy $u = (u_1, u_2)$ mean that a u-strategist displays with intensity u_1 as winner but u_2 as loser. Let $q(w, l)$ be the probability that a display of intensity w by the winner against a display of intensity l by the loser elicits submission on the part of the loser. (Thus a u-strategist wins with probability $\frac{1}{2}$ against a v-strategist, but wins and dominates only with the smaller probability $\frac{1}{2}q(u_1, v_2)$.) If a u-strategist wins, then with probability $q(u_1, v_2)$ it also dominates and its payoff is 1; but with probability $1 - q(u_1, v_2)$ it fails to dominate and its payoff is b. Let $c_w(s)$ denote the cost to a winner of displaying with intensity s. Then, conditional upon winning, a u-strategist's payoff is $q(u_1, v_2) \cdot 1 + \{1 - q(u_1, v_2)\} \cdot b - c_w(u_1)$. Likewise, conditional upon losing, a u-strategist's payoff is $q(v_1, u_2) \cdot 0 + \{1 - q(v_1, u_2)\}b - c_l(u_2)$, where $c_l(s)$ denotes the cost to a loser of displaying with intensity s. Multiplying each of the above payoffs by $\frac{1}{2}$ and adding, we find that the reward to a u-strategist in a population of v-strategists is

$$f(u,v) \;=\; \tfrac{1}{2}\big\{(1-b)q(u_1,v_2) - bq(v_1,u_2) - c_w(u_1) - c_l(u_2)\big\} + b. \tag{2}$$

We need to place conditions on the functions c_w, c_l and q. First, for c_w and c_l, it seems reasonable to suppose that $c_w(0) = 0$, $c_w'(s) > 0$, $c_w''(s) \geq 0$ (as in Mesterton-Gibbons and Sherratt, 2006) and $c_l(0) = 0$, $c_l'(s) > 0$, $c_l''(s) \geq 0$. For the sake of simplicity, we satisfy these conditions by taking

$$c_w(s) \;=\; \gamma_w \theta s, \quad c_l(s) \;=\; \gamma_l \theta s \tag{3}$$

with

$$\gamma_w \;<\; \gamma_l \tag{4}$$

throughout, where $\theta\,(>0)$ has the dimensions of INTENSITY^{-1}, so that $\gamma_w\,(>0)$ and $\gamma_l\,(>0)$ are dimensionless measures of the marginal cost of displaying for a winner and a loser, respectively.

Second, for q, the following seem reasonable: $q(\infty, l) = 1$ for any finite l, and $q(w,l) = \delta$ for all $w \leq l$ where δ is the base probability that winning will lead to dominance—a winner cannot increase its chance of converting its win into dominance unless it is displaying with at least as strong an intensity as the loser. The shorter the contest, the more likely it is that the loser will feel heavily outgunned and concede dominance; hence δ is a decreasing function of contest length T. For the sake of simplicity, we take

$$\delta \;=\; e^{-T/\mu}, \tag{5}$$

where μ is a scaling factor (the length of a contest that would reduce the probability of achieving dominance without a display from 1 to approximately 37%). We also require $\partial q/\partial w > 0$ and $\partial q/\partial l < 0$ for all $w > l$. Again for the sake of simplicity, we satisfy all conditions on q by taking

$$q(w,l) \;=\; \begin{cases} \delta + (1-\delta)\{1 - e^{-\theta(w-l)}\} & \text{if } \; w \geq l \\ \delta & \text{if } \; w < l \end{cases} \tag{6}$$

throughout. Note the asymmetry here: a display by the loser is not a second chance to win the fight. On the contrary, it is merely an attempt to reduce the probability that losing implies subordination.

3 ESS Analysis

A strategy v is an evolutionarily stable strategy or ESS in the sense of Maynard Smith (1982) if it is uniquely the best reply to itself; that is, in present circumstances, if v_1 is a winner's best reply to a loser's v_2 and v_2 is a loser's best reply to a winner's v_1. From Appendix A, if the marginal cost of displaying is so high for a winner that $\gamma_w \geq 1 - b$, then $v_1 = 0$ (i.e., not displaying) is a winner's best reply to any v_2; and likewise, if the marginal cost of displaying is so high for a loser that $\gamma_l \geq b$, then $v_2 = 0$ is a loser's best reply to any v_1. These are not interesting cases. Accordingly, we assume henceforward that $\gamma_w < 1 - b$ and $\gamma_l < b$ invariably hold. That is, we assume $\min(\rho, \zeta) > 1$, where

$$\rho \;=\; \frac{1-b}{\gamma_w}, \quad \zeta \;=\; \frac{b}{\gamma_l}. \tag{7}$$

Then, from Appendix A, and in particular from the discussion following (A8), the game defined by (2)-(6) has a unique ESS if

$$\frac{T}{\mu} \;<\; \max\Bigl\{\ln\Bigl(\frac{\rho}{\rho-1}\Bigr), \ln\Bigl(\frac{\zeta}{\zeta-1}\Bigr)\Bigr\}, \tag{8}$$

although it has no ESS if the above inequality is reversed. Subject to (8), if also

$$\frac{T}{\mu} \;<\; \min\Bigl\{\ln\Bigl(\frac{\rho}{\rho-1}\Bigr), \ln\Bigl(\frac{\zeta}{\zeta-1}\Bigr)\Bigr\}, \tag{9}$$

then from (5) and (A5) the ESS is $v = (0, 0)$: neither a winner nor a loser displays. If, on the other hand, (8) holds with (9) reversed, then one of two cases arises. If

$$\ln\Bigl(\frac{\zeta}{\zeta-1}\Bigr) \;<\; \frac{T}{\mu} \;<\; \ln\Bigl(\frac{\rho}{\rho-1}\Bigr), \tag{10}$$

then the ESS is still $v = (0, 0)$ by the remark following (A6). If

$$\ln\Bigl(\frac{\rho}{\rho-1}\Bigr) \;<\; \frac{T}{\mu} \;<\; \ln\Bigl(\frac{\zeta}{\zeta-1}\Bigr), \tag{11}$$

however, then it follows from (A6) that the ESS is given by $\theta v = (\lambda, 0)$, where

$$\lambda \;=\; \ln(\rho\{1 - e^{-T/\mu}\}). \tag{12}$$

Thus the relative magnitudes of ρ and ζ determine the ESS. For $\rho > \zeta$ or

$$\frac{\gamma_l}{\gamma_w} \;>\; \frac{b}{1-b}, \tag{13}$$

there is no ESS if $T > \mu \ln(\frac{\zeta}{\zeta-1})$; but if $T < \mu \ln(\frac{\zeta}{\zeta-1})$, then the unique ESS is given by $v = (0, 0)$ for $T < \mu \ln(\frac{\rho}{\rho-1})$ and by $\theta v = (\lambda, 0)$ for $T > \mu \ln(\frac{\rho}{\rho-1})$, with λ defined by (12). If (13) is reversed, or $\rho < \zeta$, then the unique ESS for $T < \mu \ln(\frac{\rho}{\rho-1})$ is $v = (0, 0)$; and for $T > \mu \ln(\frac{\rho}{\rho-1})$ there is no ESS.

4 Discussion

Fighting behavior and their associated signals have been the subject of extensive empirical and theoretical study (see Huntingford and Turner, 1987; Hardy and Briffa, 2012). However, much of this work has focused on the behaviors that occur before and during aggressive interactions, and relatively little is known about behaviors that occur after the outcomes have been decided (Bower, 2005). Here we have developed and explored a game-theoretic model of post-conflict signalling, seeking to identify who should tend to signal following termination of conflict, with what intensity, and the factors that shape this intensity. We have focused on the hypothesis that post-conflict signalling by the victor serves to reinforce dominance, reducing the chances that the loser will try it on again, although there may be other complementary adaptive explanations for such displays, including advertising of victory to bystanders, and non-adaptive explanations such as emotional release (Bower, 2005).

Post-conflict victory displays (Bower, 2005) have been reported in a range of organisms, including humans (Tracy and Matsumoto, 2008) and birds (Grafe and Bitz, 2004), but they have been most intensively researched in crickets (Orthoptera). Crickets often perform aggressive songs and body jerks both during and after an agonistic conflict (Alexander, 1962; Jang et al, 2008; Bertram et al, 2010). In a study of the field cricket, *Teleogryllus oceanicus*, Bailey and Stoddart (1982) proposed that if the display of a victorious male is sufficiently intense, then it may indicate to the loser that the fight is unlikely to be reversed by further combat, enabling the victor to divert its time and energy to other activities such as mating. Conversely, low signalling intensity of the winner may suggest to the loser that re-engagement could potentially produce a reversal, hence some future reward to the loser. This is precisely the situation we have attempted to model here. Indeed, Bailey and Stoddart (1982) went further and argued that the winner's post-conflict display could be used as an indication of the winner's position in a broader dominance hierarchy, showing that hierarchies constructed using an index based on post-conflict signalling correlated well with those produced by more classical methods. Intriguingly, Logue et al (2010) recently reported that contests between male field crickets *Teleogryllus oceanicus* that were unable to sing were more aggressive than interactions between males that were free to signal, supporting the view that signalling can serve to mitigate the costs of fighting in these species.

As predicted by our current model, there is now a considerable amount of evidence from the cricket literature that eventual winners tend to signal far more frequently than losers after fighting (Alexander, 1962; Bertram et al, 2010, Ting et al. in prep). One factor driving this basic result in our model (and most likely in the experiments) is our assumption that the marginal cost of signalling is lower for the winner than the loser, i.e., $\gamma_w < \gamma_l$; see (4). We consider this an entirely realistic condition given that the victor is likely to have "more left in the tank" than the vanquished (see Bower 2005, pp. 121-122 for a similar argument). Indeed, in these cases costly signals may serve as an honest indicator of how much the victor has in reserve, and thereby intimidate the opponent into submission. The "Ali shuffle" (Golus, 2006) is potentially one such example of an honest demonstration of a fighter's superiority. Analogous behaviors, which may have evolved at least in part to provide an honest signal of an individual's ability (in this case to escape predators), include stotting by Thomson's gazelles (Caro, 1986), push-up displays by anolis lizards (Leal and Rodríguez-Robles, 1997) and aerial singing by skylarks (Cresswell, 1994).

Another factor favoring the result that eventual winners tend to signal far more frequently than losers after fighting is a sufficiently low value of the parameter b, i.e., a sufficiently high reproductive advantage to dominance. Indeed it follows from the analysis in §3 above that there can be no post-conflict signalling at the ESS in our model if $\gamma_w/\gamma_l > (1-b)/b$. On the contrary, the winner victory-displays at the ESS only if (13) holds, i.e., if $\gamma_l/\gamma_w > b/(1-b)$. We have already noted that a high value of γ_l/γ_w is one way to favor this inequality, and a low value of b is clearly another.

Our model not only predicts that winners are more likely to signal than losers, but also that signalling should be more intense the longer and more intense the contest. This prediction arises simply because the ease of victory is itself a signal of dominance, a fact unlikely to be reversed through signalling either by the winner or by the loser. This phenomenon is captured in our model by making the baseline probability that winning will lead to outright dominance, i.e., δ, a decreasing function of the contest duration T; see (5). Recent studies on crickets (Bertram et al, 2010, Ting et al. in prep) support the prediction that post-conflict displays are more likely to be observed after long contests. Further evidence for this general property comes from a recent comparative study by Jang et al (2008), who examined post-conflict "dominance" displays by winners in pairwise contests of males of four different species (*Gryllus pennsylvanicus, G. rubens, G. vernalis* and *G. fultoni*, respectively). The latter two field cricket species do not fight as intensively as the former two species and, as might be anticipated, do not display as frequently following conflict, or with such vigor, as the former two species. This pattern once again suggests that only close or costly fighting selects for victory displays, although more comparative data are clearly needed.

For the model to be compared with observations in a quantitative sense, an obvious question is, how big is the reproductive value of dominance in crickets? That is, how small is b? The observation of multiple fights in a laboratory setting (Ting et al. in prep) and the very presence of clear dominance hierarchies (Bailey and Stoddart, 1982) suggest that there is a significant reproductive advantage to dominance. Suppose, in the first instance, that the reproductive advantage of dominance is high enough to ensure $b < \frac{1}{2}$. Then (13) must hold, corresponding to the shaded triangle in Figure 1, and there are two critical values of T/μ, a smaller value $\ln(\frac{\rho}{\rho-1})$, below which $v = 0$ is the ESS, and a larger value $\ln(\frac{\zeta}{\zeta-1})$, above which no ESS exists. Between these two critical values, the ESS is determined by $\theta v = (\lambda, 0)$, with λ defined by (12). We see that the intensity of the winner's display is zero until contest length reaches the first critical value, after which intensity increases

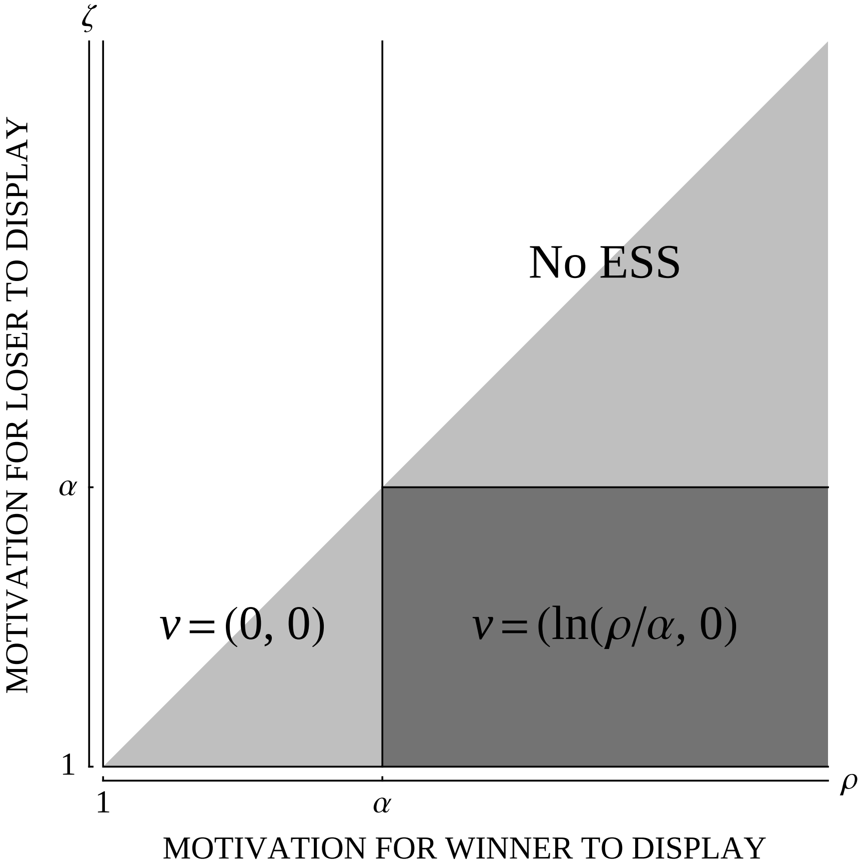

Fig. 1. ESS regions in the $\rho - \zeta$ plane for fixed values of T/μ, hence fixed values of $\alpha = 1/(1 - e^{-T/\mu})$. The shaded triangle is where $\rho = (1 - b)/\gamma_w$ exceeds $\zeta = b/\gamma_l$, i.e., where (13) holds. High reproductive advantage of dominance corresponds to low b and hence low ζ, i.e., just above the horizontal axis.

with contest length until the second critical value is reached; and the corresponding intensity of the loser's display is zero. This behavior is illustrated by Figure 2 in the limit as $\zeta \to 1$ from above, so that the second critical value recedes towards infinity.

In general, however, both critical values are finite. For contest lengths below the first critical value, neither the winner nor the loser displays at the ESS. For contest lengths between the two critical values, only the winner displays, with intensity that increases with T. For contest lengths greater than the second critical value, the ESS breaks down as described at the end of the appendix, and in such a way that a loser's optimal response will sometimes be to match the winner's display.

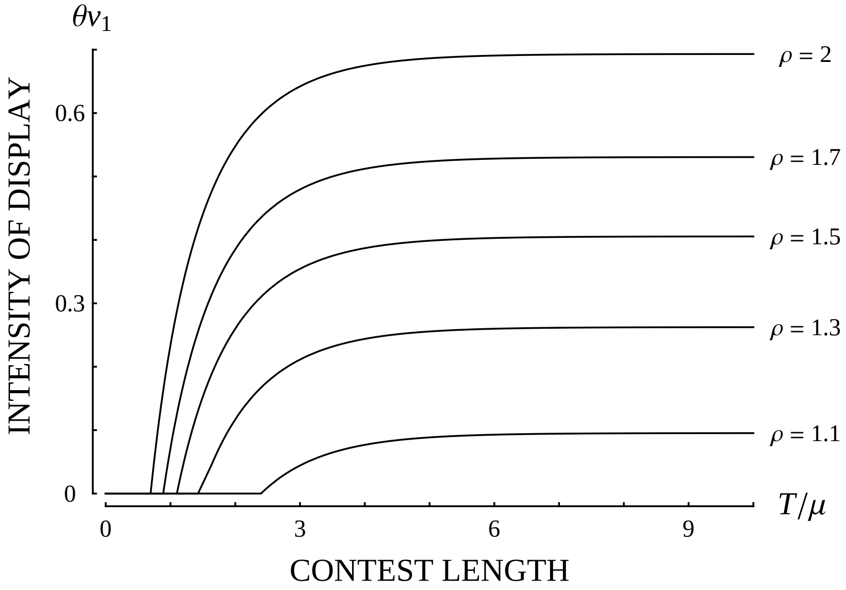

Fig. 2. Scaled intensity of winner's victory display as a function of scaled contest length for various values of the parameter $\rho = (1 - b)/\gamma_w$ (assumed to exceed 1) in the limit as $\zeta \to 1$ from above, where $\zeta = b/\gamma_l$.

Thus, according to our model, a loser should be expected to display only if the contest is so long that its length exceeds the second critical value. Those unusual biological examples in which only the loser displays (e.g., Lippold et al, 2008) may potentially be explained by some sort of subservient signal to assure dominance to the victor, thereby reducing future conflict (Bower, 2005).

There is an intriguing parallel between one of our results on victory displays and a result concerning winner effects that Mesterton-Gibbons (1999) found, several years before victory displays were first reviewed by Bower (2005). A winner effect is an increased probability of victory in a later contest following victory in an earlier contest (Rutte et al, 2006), which in Mesterton-Gibbons (1999) is mediated through increased self-perception of strength. The greater the likelihood of a later victory, the more likely it is that the earlier victory will eventually lead to dominance over the opponent. Thus a winner effect may also be regarded as an attempt to convert victory into dominance, even though there is no display. The result discovered by Mesterton-Gibbons (1999) is that there can be no winner effect unless $b < \frac{1}{2}$, where b has exactly the same interpretation as in our current model, i.e., an inverse measure of the reproductive advantage of dominance. Thus, to the extent that victory displays and winner effects can both be regarded as factors favoring dominance, such factors are most operant when $b < \frac{1}{2}$.

Finally, for the sake of tractability, we did not explicitly model the variation of strength that supports any variation of contest length observed in nature. On the contrary, we assumed that T is fixed for a theoretical population; and we obtained an evolutionarily stable response to that T, which is likewise fixed for the theoretical population. Over many such theoretical populations, each with a different T, however, there will be many different ESS responses; and in effect we have implicitly assumed that the variation of ESS with T thus engendered will reasonably approximate the variation of signal intensity with contest length observed within a single real population. Essentially this assumption—phrased more generally, that ESS variation over many theoretical populations each characterized by a different parameter value will reasonably approximate variation of behavior with respect to that parameter within a single real population—is widely adopted in the literature, although rarely made explicit, as here. Indeed essentially this assumption is made whenever a game-theoretic model predicts the dependence of an ESS on a parameter that varies within a real population, but whose variance is not accounted for by the model.

ACKNOWLEDGMENTS

We are grateful to Lauren Fitzsimmons and two anonymous reviewers for constructive feedback on earlier versions of the manuscript.

Appendix A

ESS conditions

Strategy v is a strong, global evolutionarily stable strategy or ESS in the sense of Maynard Smith (1982) if (and only if) it is uniquely the best reply to itself, in the sense that $f(v,v) > f(u,v)$ for all $u \neq v$; or, equivalently for our model, if v_1 is a winner's best reply to a loser's v_2 and v_2 is a loser's best reply to a winner's v_1.[3]

From (2) and (6) we have

$$\frac{\partial f}{\partial u_1} = \begin{cases} -\frac{1}{2}\gamma_w\theta & \text{if} \;\; u_1 < v_2 \\ \frac{1}{2}\{(1-b)(1-\delta)e^{-\theta(u_1-v_2)} - \gamma_w\}\theta & \text{if} \;\; u_1 > v_2 \end{cases} \tag{A1}$$

with $\partial^2 f/\partial u_1^2 = -\frac{1}{2}\theta^2(1-b)(1-\delta)e^{-\theta(u_1-v_2)} < 0$ for $u_1 > v_2$ but $\partial^2 f/\partial u_1^2 = 0$ for $u_1 < v_2$. So, with respect to u_1, f decreases from $u_1 = 0$ to $u_1 = v_2$. What happens next depends on the limit of $\partial f/\partial u_1$ as $u_1 \to v_2$ from above, which is $\frac{1}{2}\{(1-b)(1-\delta)-\gamma_w\}\theta$. If this quantity is not positive, then f continues to decrease, and so the maximum of f with respect to u_1 occurs at $u_1 = 0$. So a winner's best reply is $u_1 = 0$ whenever $\delta > 1 - \gamma_w/(1-b)$ (which is true in particular if $\gamma_w > 1-b$). If, on the other hand, $\delta < 1 - \gamma_w/(1-b)$, then there is a local maximum for $u_1 > v_2$ where $\partial f/\partial u_1 = 0$ or

$$\theta u_1 = \theta v_2 + \ln\biggl(\frac{(1-b)(1-\delta)}{\gamma_w}\biggr). \tag{A2}$$

The value of f at this local maximum exceeds the value at $u_1 = 0$ only if

$$\theta v_2 < \frac{(1-b)(1-\delta)}{\gamma_w} - 1 - \ln\biggl(\frac{(1-b)(1-\delta)}{\gamma_w}\biggr). \tag{A3}$$

Note that the right-hand side of (A3) is always positive (because $x - 1 - ln(x) > 0$ for all $x > 1$). In sum, a winner's best reply is $u_1 = 0$ unless $\delta < 1 - \gamma_w/(1-b)$ and (A3) holds, in which case, the best reply is given by (A2). In particular, zero is always a winner's best reply if $\gamma_w > 1-b$.

Similarly,

$$\frac{\partial f}{\partial u_2} = \begin{cases} \frac{1}{2}\{b(1-\delta)e^{\theta(u_2-v_1)} - \gamma_l\}\theta & \text{if} \;\; u_2 < v_1 \\ -\frac{1}{2}\gamma_l\theta & \text{if} \;\; u_2 > v_1 \end{cases} \tag{A4}$$

---

[3] In general, strategy v is an ESS if it does not pay a potential mutant to switch from v to any other strategy, and v need not satisfy the strong condition $f(v,v) > f(u,v)$ for all $u \neq v$. If there is at least one alternative best reply u such that $f(u,v) = f(v,v)$ but v is a better reply than u to all such u ($f(v,u) > f(u,u)$), then v is called a weak ESS. For our model, however, any ESS is a strong ESS, as is typical of continuous games (McGill and Brown, 2007, p. 408).

with $\partial^2 f/\partial u_2^2 = \frac{1}{2}\theta^2 b(1-\delta)e^{\theta(u_2-v_1)} > 0$ for $u_2 < v_1$ but $\partial^2 f/\partial u_2^2 = 0$ for $u_2 > v_1$. Note that the limit of $\partial f/\partial u_2$ as $u_2 \to v_1$ from below is $\frac{1}{2}\{b(1-\delta) - \gamma_l\}\theta$. Because $\partial^2 f/\partial u_2^2 > 0$, if the limit is negative, i.e., if $\delta > 1-\gamma_l/b$, then f decreases with respect to u_2 and has its maximum where $u_2 = 0$, so that a loser should not display. If, on the other hand, the limit is positive, i.e., $\delta < 1-\gamma_l/b$, then f at least partly increases with respect to u_2 for $u_2 < v_1$; and so the maximum of f with respect to u_2 occurs either at $u_2 = 0$ or $u_2 = v_1$, depending on which has the higher value of f. Let x_c denote the unique positive root of the equation $(1-\delta)(1-e^{-x}) = \gamma_l x/b$. Then straightforward algebra reveals that the maximum is at 0 if $\theta v_1 > x_c$ but at v_1 if $\theta v_1 < x_c$. In sum, a loser's best reply is $u_2 = 0$ unless $\delta < 1-\gamma_l/b$ and $\theta v_1 < x_c$, in which case, the best reply is v_1. Clearly, $\theta v_1 < x_c$ holds for $v_1 = 0$, so that $u_2 = 0$ is in particular the best reply to $v_1 = 0$; however, this result follows more readily directly from (A4). Also, note that zero is always a loser's best reply if $\gamma_l > b$.

For $v = (v_1, v_2)$ to be an ESS it must be a best reply to itself, i.e., we require v_1 to be a winner's best reply to the loser's v_2 at the same time as v_2 is a loser's best reply to the winner's v_1. If

$$\delta \;>\; \max\left(1 - \frac{\gamma_w}{1-b}, 1 - \frac{\gamma_l}{b}\right) \tag{A5}$$

then the unique ESS is $v = (0,0)$, because it follows from the discussion after (A1) that $v_1 = 0$ is the best reply to any v_2, and hence to $v_2 = 0$; and from the discussion after (A4) that $v_2 = 0$ is the best reply to any v_1, and hence to $v_1 = 0$. If

$$1 - \frac{\gamma_w}{1-b} \;<\; \delta \;<\; 1 - \frac{\gamma_l}{b} \tag{A6}$$

then $v = (0,0)$ is still the ESS by the above discussion and the remark at the end of the preceding paragraph. If instead

$$1 - \frac{\gamma_l}{b} \;<\; \delta \;<\; 1 - \frac{\gamma_w}{1-b} \tag{A7}$$

then the ESS is given by $\theta v = (\lambda, 0)$ where λ is defined by (12), because it follows from (A2) that $v_1 = \lambda/\theta$ is the best reply to $v_2 = 0$; and $v_2 = 0$ is still the best reply to any v_1, and hence to $v_1 = 0$.

If, on the other hand,

$$\delta \;<\; \min\left(1 - \frac{\gamma_w}{1-b}, 1 - \frac{\gamma_l}{b}\right) \tag{A8}$$

then an ESS does not exist. Consider a population in which a winner displays with small positive intensity v_1. Then $\theta v_1 < x_c$; and, from the discussion following (A4), a loser's best reply is to match the display. From (A2), a winner's best reply is now to increase the intensity of its display, because (A3) invariably holds; and a loser's best reply in turn is again to match the display. Continuing in this manner, we observe an "arms race" of increasing display intensity, until either (A3) or $\theta v_1 < x_c$ is violated. If the former, then a winner's best reply is not to display, which a loser matches, so that it pays for a winner to display at higher intensity; if the latter, then a loser's best reply becomes no display, but now a winner's best reply is to display with intensity λ/θ. Either way, the unstable cycle continues ad infinitum. The study of victory displays is still in its infancy, and researchers are still trying to characterize when it occurs and with what frequency. Therefore, there is no study into its temporal dynamics (within or between generations). Furthermore, full analysis of the evolutionary dynamics when no ESS exists is beyond the scope of this paper. Nevertheless, we broach this issue in Appendix B.

Appendix B

What happens when there is no ESS?

In this appendix we remark on why, when no ESS exists, the evolutionary dynamics require a more sophisticated approach than the one we have taken in this paper and cannot readily be addressed by the standard framework of discrete evolutionary games with replicator dynamics (e.g., Hofbauer and Sigmund, 1998; Cressman, 2003). To make our point as expeditiously as possible, we explore circumstances in which $b < \frac{1}{2}$ and hence $\rho > \zeta$ but T/μ exceeds the second critical value $\ln(\frac{\zeta}{\zeta-1})$ of §3 (corresponding to the shaded triangle within the no-ESS region of Figure 1).

Accordingly, consider a mixture of three strategies that appear to evoke the discussion towards the end of Appendix A, namely, a non-signalling strategy, denoted by N or Strategy 1; the ESS signalling strategy for the dark shaded rectangle of Figure 1, denoted by S or Strategy 2; and a matching strategy, denoted by M or strategy 3, which displays with the ESS intensity corresponding to $\ln(\frac{\rho}{\rho-1}) < T/\mu < \ln(\frac{\zeta}{\zeta-1})$ after winning, but matches the winner's display after losing. From the viewpoint of a focal u-strategist against a v-strategist, these three strategies are defined, respectively, by $u = (0,0)$ for N; $u = (\lambda/\theta, 0)$ for S; and $u = (\lambda/\theta, v_1)$ for M. Let the proportions of N, S and M be x_1, x_2 and x_3, respectively (so that $x_1 + x_2 + x_3 = 1$); and let a_{ij} be the reward to strategy i against strategy j (for $1 \leq i,j \leq 3$). Then from (2), (6) and (12) we have $a_{11} = \frac{1}{2}\{(1-2b)q(0,0) - c_w(0) - c_l(0)\} + b = \frac{1}{2}\delta + (1-\delta)b$, $a_{12} = \frac{1}{2}\{(1-b)q(0,0) - bq(\lambda/\theta, 0) - c_w(0) - c_l(0)\} + b = \frac{1}{2}\{(1-b)\delta + b(1+1/\rho)\}$, and so on, yielding the reward matrix

$$A = \begin{bmatrix} \frac{1}{2}\delta + (1-\delta)b & \frac{\rho(1-b)\delta+(\rho+1)b}{2\rho} & a_{12} \\ \frac{\rho-1+b+(1-\delta)\rho b}{2\rho} - \frac{1}{2}\gamma_w \ln(\rho/\alpha) & \frac{\rho-1+2b}{2\rho} - \frac{1}{2}\gamma_w \ln(\rho/\alpha) & a_{12} - \frac{1}{2}\gamma_w \ln(\rho/\alpha) \\ a_{21} & a_{21} - \frac{1}{2}\gamma_l \ln(\rho/\alpha) & a_{11} - \frac{1}{2}(\gamma_w + \gamma_l) \ln(\rho/\alpha) \end{bmatrix} \qquad \text{(B1)}$$

where $\alpha = 1/(1-\delta)$, as in Figure 2.

Because $\gamma_w \rho = 1 - b$ and $\rho/\alpha = \rho(1-\delta) > 1$ by assumption, $a_{11} - a_{21} = \frac{1}{2}\gamma_w\{1 - \rho/\alpha + \ln(\rho/\alpha)\}$ must be negative (because $1 - x + \ln(x) < 0$ for all $x > 1$). Thus $a_{11} < a_{21}$, and Strategy 1 is not an ESS. However, because $a_{11} - a_{21} + a_{22} - a_{12} = 0$, it also follows that $a_{22} > a_{12}$. So if Strategy 2 is also not an ESS, then it must be Strategy 3 that invades. But $a_{22} - a_{32} = \frac{1}{2}\gamma_l\{(1/\rho - 1/\alpha)\zeta + \ln(\rho/\alpha)\}$ may have either sign, and in particular will always be positive for sufficiently small ζ, that is, for ζ sufficiently close to α ($\zeta > \alpha$ having been assumed). Furthermore, if we suppose that the point (ρ, ζ) in Figure 1 has migrated from the signalling ESS region (dark shaded rectangle) into the no-ESS region (shaded triangle immediately above) because environmental pressures have increased the value the value of ζ (by decreasing γ_l), then it is precisely such sufficiently small values of ζ that are relevant. Thus S will often be an ESS of the discrete game defined by the matrix A even though it is no longer an ESS of the continuous game described in the main body of our paper.

The upshot is that replicator dynamics cannot readily be used to describe what happens when S is not an ESS of our continuous game; the dynamics described verbally towards the end of Appendix A are not adequately reflected by a mix of M, N and S. They require a much more sophisticated approach, and we leave the matter open for future work.

Nevertheless, let us suppose that ζ is indeed large enough for M to invade S, that is,

$$\zeta > \frac{\rho\alpha}{\rho - \alpha}\ln(\rho/\alpha). \tag{B2}$$

Then $a_{22} - a_{32} < 0$, and because $a_{33} - a_{23} + a_{22} - a_{32} = 0$, we must have $a_{33} - a_{23} > 0$. That is, from (B1), $a_{33} > a_{13} - \frac{1}{2}\gamma_w \ln(\rho/\alpha)$. Thus $a_{33} > \max(a_{13}, a_{23})$ will hold for sufficiently small γ_w, making M the unique ESS of the discrete game defined by A. Otherwise (i.e., for larger values of γ_w), $a_{32} > a_{22}$, $a_{21} = a_{31} > a_{11}$ and $a_{13} > a_{33}$ will hold simultaneously, so that M can invade S, S or M can invade N and N can invade M. In these circumstances, the population will eventually reach a polymorphism of N and M at which $x_1 = (a_{13} - a_{33})/(a_{13} - a_{33} + a_{31} - a_{11})$, $x_2 = 0$ and $x_3 = 1 - x_1$. All of these results have been verified by numerical integration, for relevant parameter values, of the replicator equations $\dot{x}_i = x_i\{(Ax)_i - x \cdot Ax\}, i = 1, 2, 3$, where $x = (x_1, x_2, x_3)$ and an overdot denotes differentiation with respect to time (see, e.g., Hofbauer and Sigmund, 1998, p. 68).

REFERENCES

Alexander RD (1962) Evolutionary change in cricket acoustical communication. Evolution 16:443–467

Bailey WJ, Stoddart JA (1982) A method for the construction of dominance hierarchies in the field cricket *Teleogryllus oceanicus* (Le Guillou). Animal Behaviour 30:216–220

Bertram SM, Rook VLM, Fitzsimmons LP (2010) Strutting their stuff: victory displays in the spring field cricket, *Gryllus veletis*. Behaviour 147:1249–1266

Bower JL (2005) The occurrence and function of victory displays within communication networks. In: McGregor P (ed) Animal Communication Networks, Cambridge University Press, Cambridge, pp 114–126

Caro TM (1986) The functions of stotting in Thomson's gazelles: some tests of the predictions. Animal Behaviour 34:663–684

Cressman R (2003) Evolutionary Dynamics and Extensive Form Games. MIT Press, Cambridge, Massachusetts

Cresswell W (1994) Song as a pursuit-deterrent signal, and its occurrence relative to other anti-predation behaviours of skylark (*Alauda arvensis*) on attack by merlins (*Falco columbarius*). Behavioral Ecology and Sociobiology 34:217–223

Golus C (2006) Muhammad Ali. Twenty-First Century Books, Minneapolis, Minnesota

Grafe T, Bitz JH (2004) An acoustic postconflict display in the duetting tropical boubou (*Laniarius aethiopicus*): a signal of victory? BMC Ecology 4(1), http://www.biomedcentral.com/1472-6785/4/1

Hardy ICW, Briffa M (eds) (2012) Animal Contests. Cambridge University Press, Cambridge, in press

Hofbauer J, Sigmund K (1998) Evolutionary Games and Population Dynamics. Cambridge University Press, Cambridge, UK

Huntingford F, Turner AK (1987) Animal Conflict. Chapman & Hall, London

Jang Y, Gerhardt HC, Choe JC (2008) A comparative study of aggressiveness in eastern North American field cricket species (genus *Gryllus*). Behavioral Ecology and Sociobiology 62:1397–1407

Leal M, Rodríguez-Robles JA (1997) Signalling displays during predator-prey interactions in a puerto rican anole, *Anolis cristatellus*. Animal Behaviour 54:1147–1154

Lippold S, Fitzsimmons LP, Foote JR, Ratcliffe LM, Mennill DJ (2008) Post-contest behaviour in black-capped chickadees (*Poecile atricapillus*): loser displays, not victory displays, follow asymmetrical countersinging exchanges. Acta Ethologica 11:67–72

Logue DM, Abiola IO, Rains D, Bailey NW, Zuk M, Cade WH (2010) Does signalling mitigate the cost of agonistic interactions? a test in a cricket that has lost its song. Proceedings of the Royal Society of London B 277:2571–2575, DOI 10.1098/rspb.2010.0421

Maynard Smith J (1982) Evolution and the Theory of Games. Cambridge University Press, Cambridge

McGill BJ, Brown JS (2007) Evolutionary game theory and adaptive dynamics of continuous traits. Annual Review of Ecology and Systematics 38:403–435

Mesterton-Gibbons M (1999) On the evolution of pure winner and loser effects: a game-theoretic model. Bulletin of Mathematical Biology 61:1151–1186

Mesterton-Gibbons M, Sherratt TN (2006) Victory displays: a game-theoretic analysis. Behavioral Ecology 17:597–605, DOI doi:10.1093/beheco/ark008

Rutte C, Taborsky M, Brinkhof MWG (2006) What sets the odds of winning and losing? Trends in Ecology and Evolution 21:16–21

Tracy JL, Matsumoto D (2008) The spontaneous expression of pride and shame: evidence for biologically innate nonverbal displays. Proceedings of the National Academy of Sciences USA 105:11,655–11,660