

**This is an electronic reprint of the original article.  
This reprint *may differ* from the original in pagination and typographic detail.**

**Author(s):** Mesterton-Gibbons, Mike; Heap, Stephen

**Title:** Variation between Self- and Mutual Assessment in Animal Contests

**Year:** 2014

**Version:**

**Please cite the original version:**

Mesterton-Gibbons, M., & Heap, S. (2014). Variation between Self- and Mutual Assessment in Animal Contests. *The American Naturalist*, 183(2), 199-213.  
<https://doi.org/10.1086/674443>

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

# Variation between Self- and Mutual Assessment in Animal Contests

Mike Mesterton-Gibbons<sup>1,\*</sup> and Stephen M. Heap<sup>2</sup>

1. Department of Mathematics, Florida State University, Tallahassee, Florida 32306; 2. Centre of Excellence in Biological Interactions, Department of Biological and Environmental Science, University of Jyväskylä, Jyväskylä 40014, Finland

Submitted January 31, 2013; Accepted September 19, 2013; Electronically published December 24, 2013

Online enhancement: appendixes.

---

**ABSTRACT:** Limited resources lead animals into conflicts of interest, which are resolved when an individual withdraws from a direct contest. Current theory suggests that the decision to withdraw can be based on a threshold derived from an individual's own state (self-assessment) or on a comparison between their own state and their opponent's (mutual assessment). The observed variation between these assessment strategies in nature does not conform to theory. Thus, we require theoretical developments that explain the functional significance of different assessment strategies. We consider a hawk-dove game with two discrete classes that differ in fighting ability, in which the players strategically decide on their investment toward mutual assessment. Analysis of the model indicates that there are simultaneous trade-offs relating to assessment strategies. First, weak individuals in a population must decide on whether to acquire information about their opponents at the cost of providing opponents with information about themselves. Secondly, all individuals must decide between investing in mutual assessment and being persistent in contests. Our analysis suggests that the potential for individuals to make errors during contests and differences in the consequences of sharing information within a population may serve as fundamental concepts for explaining variation in assessment strategy.

*Keywords:* contest, self-assessment, mutual assessment, signaling, information, strategic error.

---

## Introduction

Animals can have conflicts of interest over access to resources that can be decided by a contest (Maynard Smith 1974; Kokko 2013). Individuals engaged in a contest are faced with a trade-off between gaining the benefits of resource access and minimizing the cost of interactions with rivals, which is resolved by a strategic decision to withdraw from the interaction (Arnott and Elwood 2008, 2009). The decision to withdraw from a contest, rather than risk further costs by persistence, can be made more effective by

responding to information that can predict the outcome (Giraldeau 1997; Dall et al. 2005).

One potential source of information is the resource holding potential (RHP) of each individual: each contestant's ability to physically overcome its opponent or pay the costs of contesting resource ownership (Parker 1974; Parker and Rubenstein 1981). An individual's RHP is dependent on the particulars of resource competition for its species; common factors include size, mass, and resource stores, which provide the ability to inflict damage, suffer blows, and metabolize energy (Clutton-Brock and Albon 1979; Hsu et al. 2006; Arnott and Elwood 2009). RHP can also correlate with morphological structures or signals (e.g., Stuart-Fox et al. 2006), although the links between these traits and fighting ability is not always clear. In any case, factors that influence RHP are typically able to be measured or estimated by researchers and are thus potentially assessable by the species in question as well. In this regard, studies have supported a role for RHP in contests and suggest that animals can improve their decision making in contests by perceiving and responding to information on RHP asymmetries (Arnott and Elwood 2009). Theoretical models have been used to consider optimal responses to this information, yet they typically assume that individuals are either completely ignorant or in complete possession of information on RHP asymmetries (Crowley 2000). By contrast, little is known about how individuals may behave when they have only partial information on the RHP asymmetry, especially when they have the option to improve on this information via a costly investment.

There are at least two sources of information (sensu Wagner and Danchin 2010) on RHP: private information on an individual's own RHP and public information regarding the RHP of an opponent. Furthermore, there are at least two sources of public information on opponent RHP. Individuals may first have access to information on the distribution of RHP within the population (general

\* Corresponding author; e-mail: mesterto@math.fsu.edu.

information) and second to cues or signals that indicate the RHP of a given opponent (specific information). Provided that an individual has some private and public information, it can compare its own RHP to its opponent's RHP (based on either general or specific information) in order to assess whether there is an RHP asymmetry that can influence the outcome of the contest. That is, an individual can assess the likely payoffs involved with fighting by using general information to compare its own RHP to the expected frequencies of opponents with different RHP. Individuals can then improve on this estimate by gathering specific information from their current opponent and can benefit by basing their decision of whether to withdraw or persist in an encounter on this assessment (Parker and Rubenstein 1981).

Whether general or specific information is used in an RHP comparison can determine the process of RHP assessment in a contest and the withdrawal strategy that results from this assessment. Theory suggests that individuals can respond to general information by assessing their own RHP—a self-assessment strategy (Mesterton-Gibbons et al. 1996; Payne and Pagel 1996, 1997; Payne 1998). For instance, individuals may persist or withdraw from contests according to the probability of encountering an individual of lower or higher RHP, respectively. More specifically, models that consider the structure of contests suggest that animals can be involved in a war of attrition in which they persist until reaching a personal threshold that is determined by their own RHP and the distribution of RHP in the population (Mesterton-Gibbons et al. 1996). In contrast, access to specific information allows animals to directly compare their own RHP against the RHP of their current opponent and withdraw only if they are certain that they are outmatched. Contests based on this mechanism use a mutual assessment strategy (Parker and Rubenstein 1981; Hammerstein and Parker 1982; Enquist and Leimar 1983; Enquist et al. 1990).

Although animals can benefit from responding to the RHP asymmetry, the information that reflects the asymmetry is inherently uncertain and not completely reliable. Thus, there is always the potential for an individual to make a strategic error in its decision of whether or not to withdraw (Parker and Rubenstein 1981; Hammerstein and Parker 1982; Enquist and Leimar 1983). That is, individuals may persist when they would benefit more from withdrawing or withdraw when it would have been better to persist. These errors can result in unnecessary costs incurred through fruitless persistence and lost opportunities to access a resource, respectively. The potential for these errors to be made has a strong influence on adaptive withdrawal strategies (Crowley 2000). Theoretical and empirical evidence has suggested that it can be adaptive to reduce the risk of strategic error by gathering specific information

(Enquist and Leimar 1983), which led to the proposition that mutual assessment processes predominate in animal contests.

Recent work has brought the validity of empirical studies that detect mutual assessment into question (Bridge et al. 2000; Taylor et al. 2001; Taylor and Elwood 2003; Briffa and Elwood 2009). Such studies suggest that assessment strategy is variable between and within species (e.g., Hsu et al. 2008; Prenter et al. 2008; Kasumovic et al. 2010). Furthermore, the intensified scrutiny of recent studies has uncovered a variety of assessment modes that were not predicted by theory (Arnott and Elwood 2009), and there are multiple cases of inconclusive results concerning the assessment strategies being used within a population (Jennings et al. 2005; Morrell et al. 2005; Stuart-Fox 2006). Therefore, it is clear that a functional understanding of the basis of withdrawal decisions in contests is lacking.

It has been suggested that variation in assessment strategy occurs as the result of a trade-off between the costs of reliably gathering specific information and the benefits of shorter and less intense contests provided by acting on this information (Taylor and Elwood 2003; Prenter et al. 2006; Elias et al. 2008; Arnott and Elwood 2009). However, this hypothesis has not been empirically tested or formally examined. We thus developed a model in an effort to initiate a general theory for explaining variation in assessment strategy. As current models of contest behavior do not consider variation in assessment strategy, we used the argument for partial mutual assessment as the theoretical basis for our model (Prenter et al. 2006). This hypothesis suggests that individuals have reliable information on their own RHP but can gather some degree of information on a specific opponent's RHP. Although Prenter et al.'s (2006) verbal model of variable assessment strategy has yet to receive any formal validation, it does have some empirical support (Kelly 2006; Smallegange et al. 2007; Elias et al. 2008). The model presented in this article formalizes the concept of variable assessment to understand how the costs and benefits of gathering specific information determine the adaptive use of self- or mutual assessment strategies. Furthermore, by considering information that varies in its cost and ability to predict outcomes, it is able to account for the effect of uncertainty and the potential to make strategic errors during contests.

### Model Description

A summary of the model is given in figure 1, and its parameters—which we introduce in this and the following section—are listed in table 1 for ease of reference. We consider a population of individuals in which access to a fitness-enhancing resource of value  $V$  is determined through direct interactions among pairs (players 1 and 2).

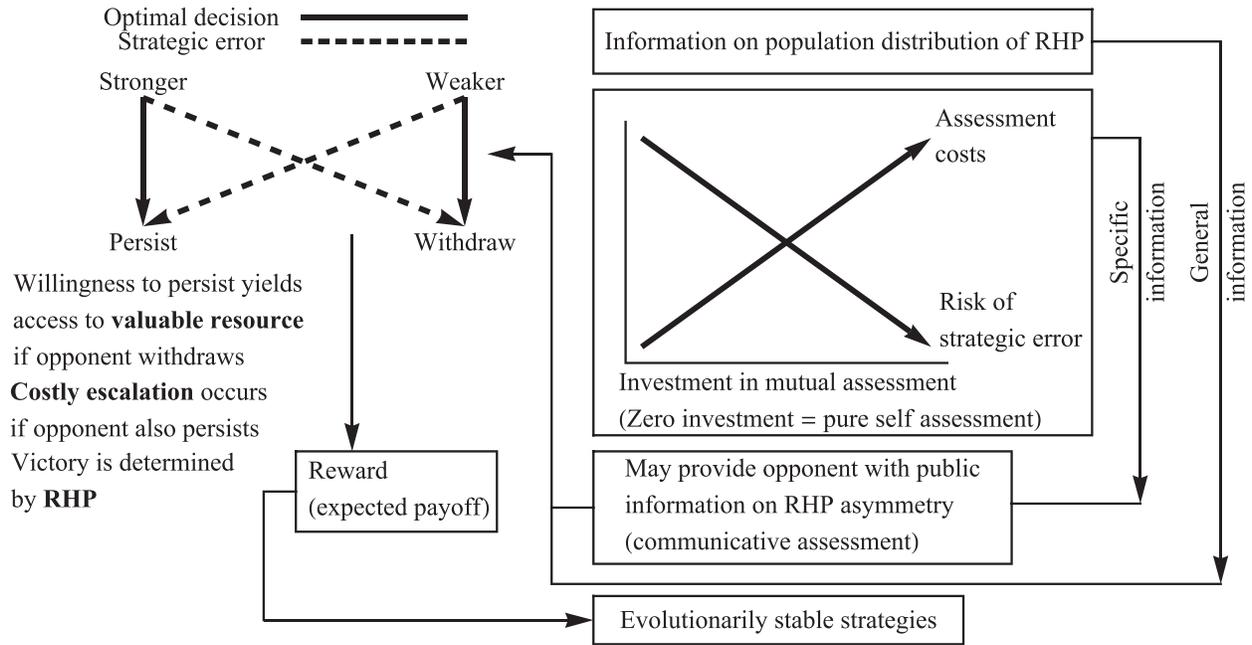


Figure 1: Pictorial representation of model's major components

This resource could represent anything that improves fitness, including food, mating opportunities, or an advantageous social position. We assume that the value of the resource is symmetrical in that each individual would receive equivalent benefits on winning a contest. Within these interactions, individuals use a withdrawal strategy that corresponds to a hawk-dove game (Maynard Smith and Parker 1976), in which they have the option to withdraw from the contest by playing dove or persist by playing hawk. The two individuals are assumed to make their decisions simultaneously, and the result of the interaction determines the fitness consequences for each individual. If both individuals play hawk, then the interaction escalates: both individuals suffer a fight cost  $C$ , and one individual gains access to the resource. Fight costs can include the time and energy devoted to the contest (Clutton-Brock and Albon 1979; Hack 1997a,1997b; Briffa and Sneddon 2007), the buildup of metabolic wastes (Briffa and Sneddon 2007; Copeland et al. 2011), increased predation risk (Jakobsson et al. 1995), physical damage, or the risk of fatality (Enquist et al. 1990; Murray 1990; Haley 1994; Payne 1998; West et al. 2001). We assume that the winner of any fight is determined solely by RHP. For simplicity, we assume that individuals can either be weak, W, or strong, S, with strong individuals always winning against weak individuals and evenly matched opponents having an equal chance of winning. If only one individual plays hawk, then it will gain access to the resource without

a fight; and if both individuals play dove, then the resource is assumed to be shared or randomly allocated (i.e., payoff =  $1/2 V$ ).

We assume that the proportions of individuals in the population having weak and strong RHP are  $\lambda$  and  $1 - \lambda$ , respectively, with at least some individuals of each type, that is,

$$0 < \lambda < 1. \quad (1)$$

Additionally, we assume that individuals have perfect knowledge of their own RHP, but not the RHP of their opponent.

From the perspective of a given individual, the world has four potential RHP combinations, which we have labeled  $i = 1-4$ , namely, WW (weak self, weak opponent;  $i = 1$ ), WS (weak self, strong opponent;  $i = 2$ ), SW (strong self, weak opponent;  $i = 3$ ), and SS (strong self, strong opponent;  $i = 4$ ). The probability of any one of these states occurring is simply a function of the distribution of RHP in the population; specifically, if  $p_i$  denotes the probability of state  $i$ , then

$$p_1 = \lambda^2, p_2 = \lambda(1 - \lambda) = p_3, p_4 = (1 - \lambda)^2. \quad (2)$$

The RHP combination that an individual actually perceives is determined by its strategic investment toward acquiring information about its current opponent (i.e., self- or mutual assessment). Because the information in equation (2) is a facet of the RHP distribution in the population, it

**Table 1:** Model parameters

Parameter	Meaning	Range	Key?
$V$	Value of fitness-enhancing resource	$V > 0$	
$C$	Cost of (winning or losing) a fight	$C > 0$	
$\lambda$	Proportion of individuals with weak (W) RHP in population	$0 < \lambda < 1$	Yes
$r$	Receptivity, i.e., the degree to which investment toward mutual assessment communicates information to the opponent as well as to the investor (with assessment being noncommunicative if $r = 0$ )	$0 \leq r \leq 1$	Yes
$\alpha$	Efficacy of assessment, i.e., the capacity for effort invested toward mutual assessment to correctly determine an opponent's RHP	$\alpha > 0$	
$\Gamma$	Cost per unit effort of assessment	$\Gamma > 0$	
$\beta = V/C$	Relative resource value, i.e., value of resource relative to cost of escalation	$\beta > 0$	Yes
$\Gamma/\alpha$	Effective assessment cost, i.e., the assessment cost of the effort required to achieve a specific relative reduction in the probability of perception error, as described at the end of app. A, available online	$\Gamma/\alpha > 0$	
$\gamma = \Gamma/\alpha C$	Relative assessment cost, i.e., effective assessment cost $\Gamma/\alpha$ relative to cost of escalation $C$	$\gamma > 0$	Yes

Note: In the first instance, the model is defined in terms of the six parameters in the first six rows of the table. However,  $\Gamma$  and  $\alpha$  appear only in the combination  $\Gamma/\alpha$ , which reduces the number to five; and scaling fitness with respect to escalation cost  $C$  further reduces the number to four. These four dimensionless parameters are the only ones on which the ESS depends and in this sense are the key parameters, as indicated by the right-most column. RHP = resource holding potential.

is public, and so all individuals are assumed to be initially able to perceive their opponent as weak with probability  $\lambda$  and strong with probability  $1 - \lambda$ . Thus, individuals have both perfect private (own RHP) and general (distribution of RHP in the population) information, allowing them to assess the average RHP asymmetry that they will encounter. This information could have been gathered over generations, in the form of genes, or over an individual's lifetime as a result of learning from personal experience (Whitehouse 1997; Maynard Smith 2000; Dall et al. 2005; Hsu et al. 2006; Arnott and Elwood 2009; Wagner and Danchin 2010). By having only two discrete RHP classes, making RHP a perfect predictor of fight outcome and making the additional assumption that RHP is reliably communicated (i.e., with no possibility for bluff), we can keep the model as simple as possible. Furthermore, coupled with the assumption that private and general information is perfect, we can focus attention on the importance of strategic error as a result of incorrect assessment, including the influence of these errors on individuals with different RHP, with as few complicating factors as possible.

In addition to using purely general information, individuals have the option to improve their information on a specific opponent's RHP through mutual assessment. Individuals strategically decide on how much effort to invest in assessing their opponent; this effort is  $u_1$  for player 1 and  $v_1$  for player 2, where  $u_1, v_1 \geq 0$ . Furthermore, we

distinguish between communicative and noncommunicative assessment behaviors, as assessment is not necessarily passive or unilateral and can depend on the actions taken by each contestant (Hack 1997*b*). Under noncommunicative assessment, assessing an opponent provides an individual with information, but provides no information to its opponent (e.g., a visual assessment). In contrast, individuals either advertently or inadvertently provide their opponent with information on the RHP asymmetry if assessment is communicative (e.g., a tug of war or comparison of signal strength). We establish the degree of communicative assessment by assuming that a given investment toward mutual assessment is as effective at yielding information to the opponent as to the investor according to how receptive individuals are to each other's behavior. We represent receptivity by  $r$ , with assessment being noncommunicative when  $r = 0$  and communicative when  $r > 0$ . Accordingly, let  $E$  denote the total effective effort directed at judging the other's RHP correctly. Then for player 1 we set

$$E = X = u_1 + rv_1, \quad (3)$$

while for player 2 we set

$$E = Y = ru_1 + v_1. \quad (4)$$

Let  $\phi_w(E)$  be the probability that an animal perceives the other contestant correctly if the other animal is weak,

and let  $\phi_s(E)$  be the probability that an animal perceives the other contestant correctly if the other animal is strong. Then, from above, it is reasonable to assume that

$$\phi_w(0) = \lambda, \phi_s(0) = 1 - \lambda, \quad (5)$$

and  $\phi_w(\infty) = \phi_s(\infty) = 1$  (i.e., infinite effort is required for perfect perception). To model the nature of how assessment accuracy changes in response to increasing effort (independently of opponent RHP), we first consider the potential for being incorrect using mutual assessment (when  $E > 0$ ) relative to the potential for being incorrect using self-assessment (when  $E = 0$ ). Specifically, the probability of making an assessment error decreases with total effort ( $E$ ) according to

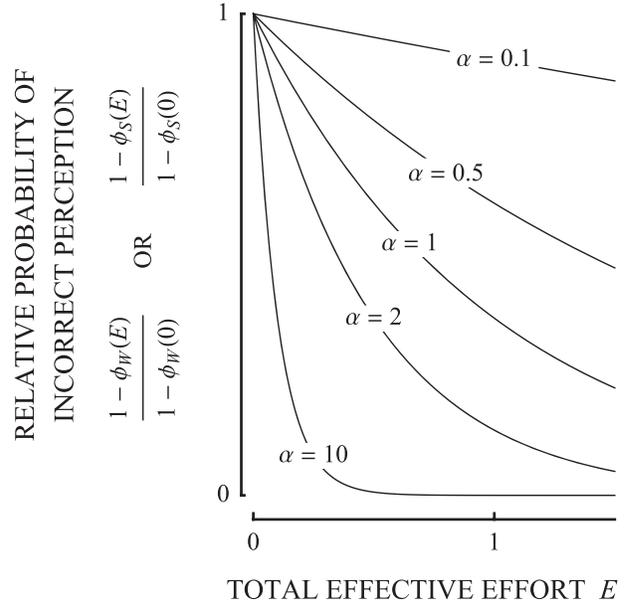
$$\frac{1 - \phi_w(E)}{1 - \phi_w(0)} = \frac{1 - \phi_s(E)}{1 - \phi_s(0)} = e^{-\alpha E}. \quad (6)$$

Here, the probability of making an incorrect perception (given effort toward assessment) relative to the initial probability of being incorrect (i.e., pure self-assessment) is determined by an exponential rate parameter,  $\alpha$  (fig. 2). The rate parameter  $\alpha$  represents the efficacy of assessment (i.e., the capacity for effort invested toward mutual assessment to correctly determine the opponent's RHP). This could reflect the sensory and cognitive abilities of the assessor, the reliability of signaling, or previous experience with the opponent. Thus, from equations (5) and (6), the respective probabilities of perceiving a weak or strong opponent correctly increase with total effective effort according to

$$\begin{aligned} \phi_w(E) &= 1 - \{1 - \phi_w(0)\}e^{-\alpha E} = 1 - (1 - \lambda)e^{-\alpha E}, \\ \phi_s(E) &= 1 - \{1 - \phi_s(0)\}e^{-\alpha E} = 1 - \lambda e^{-\alpha E}. \end{aligned} \quad (7)$$

These equations simply state that the probability of being correct ( $\phi_w(E)$  or  $\phi_s(E)$ ) can be calculated as the complementary probability of the probability of making an incorrect assessment using only self-assessment ( $1 - \phi_w(0)$  or  $1 - \phi_s(0)$ ), discounted by the effect of mutual assessment.

Additionally, mutual assessment is not free: it incurs a cost  $\Gamma$  per unit of effort  $u_1$  or  $v_1$ . Such assessment costs can arise from a number of sources, although few studies have explicitly considered them. For instance, individuals may need to produce signals that are costly in time or energy in order to receive honest information from their opponent (Dawkins and Guilford 1991), with the cost of signaling increasing as the information content of the signal increases (Enquist et al. 1990; Hack 1997b). Additionally, receiving information from an opponent may expose an individual to predation risk (Dawkins and Guilford 1991) or may involve being damaged by an opponent's



**Figure 2:** Probability of an incorrect perception relative to the initial probability of an incorrect perception as a function of total effort toward assessing an opponent's resource holding potential (RHP), for various values of the efficacy  $\alpha$ . This parameter measures the capacity for effort invested toward mutual assessment to correctly determine the opponent's RHP, in the sense that a small investment greatly reduces the probability of error when  $\alpha$  is high but has comparatively little effect when  $\alpha$  is low.

attack (Enquist and Leimar 1983; Payne 1998). There may also be costs associated with the development of cognitive or perceptive abilities required for making assessments (Taylor and Elwood 2003; Prenter et al. 2006; Elias et al. 2008; Arnott and Elwood 2009; Elwood and Arnott 2012).

For each of the four possible actual states  $i$  in the model, there are four possible perceptual states, which we label  $j = 1-4$ : CC (both animals perceive their opponent's RHP correctly;  $j = 1$ ), CI (only player 1 perceives correctly;  $j = 2$ ), IC (only player 2 perceives correctly;  $j = 3$ ), and II (neither animal perceives the other's RHP correctly;  $j = 4$ ).

Each contestant is assumed to respond to its perception of the RHP combination by playing hawk if it perceives SW and dove if it perceives WS. That is, weak individuals will always withdraw from an interaction if they perceive themselves to be outmatched, and strong individuals will always be willing to press their perceived advantage. However, individuals that perceive themselves to be matched with their opponent have a range of strategic withdrawal options available to them, because individuals may escalate with any probability between 0 (pure dove) and 1 (pure hawk). Individuals that perceive the RHP combination as WW play hawk with strategic probability  $u_2$  for player 1 and  $v_2$  for player 2, hence, dove with probability  $1 - u_2$

**Table 2:** Evolutionarily stable strategies (ESSs) with self- or mutual assessment classified according to the strategic withdrawal components of the ESS

Case	$v_2^*$	$v_3^*$	Self-assessment type	Mutual-assessment type	Character
a	1	1	IA: strong (sect. B.1.1)	IIa: strong (sect. B.2.1)	Daring
b	•	1	IB: weak (sect. B.1.9)	IIb: weak (sect. B.2.9)	Daring
c	0	1	IC: strong (sect. B.1.3)	No such ESS (sect. B.2.3)	Conventional
d	0	•	ID: weak (sect. B.1.5)	II d: weak (sect. B.2.5)	Careful
e	0	0	IE: strong (sect. B.1.2)	II e: strong (sect. B.2.2)	Careful
f	•	•	No such ESS (sect. B.1.6)	II f: weak (sect. B.2.6)	Mixed
g	•	0	No such ESS (sect. B.1.8)	II g: weak (sect. B.2.8)	Anomalous

Note:  $v^* = (v_1^*, v_2^*, v_3^*)$  with  $v_1^* = 0$  for self-assessment and  $v_1^* > 0$  for mutual assessment;  $v_2^*$  and  $v_3^*$  are the probabilities of persisting when perceiving WW and SS, respectively; and a dot denotes a positive number less than 1. Also, “strong” or “weak” refers to type of evolutionarily stable strategy, as opposed to resource holding potential type: mutant strategies are always selected against at a strong ESS, whereas mutant strategies can spread initially before being selected against at a weak ESS. The sections named in the table are available online.

or  $1 - v_2$ . Those that perceive SS play hawk with strategic probability  $u_3$  for player 1 and  $v_3$  for player 2, hence, dove with probability  $1 - u_3$  or  $1 - v_3$ . An individual that has an incorrect perception of the RHP combination makes a strategic error in terms of its withdrawal strategy: responding to a weak opponent as if it were strong, or to a strong opponent as if it were weak.

From all of the above, we can assign a behavioral strategy to each individual in the population. Specifically, the strategy consists of a three-dimensional vector comprising (i) how much to invest in acquiring specific information and how often to play hawk on perceiving a matched opponent when (ii) weak or (iii) strong. This strategy is formally represented by  $u = (u_1, u_2, u_3)$  for player 1 and  $v = (v_1, v_2, v_3)$  for player 2. The components of the strategy are constrained according to  $u_i, v_i \geq 0$ , and  $0 \leq u_i, v_i \leq 1$  for  $i = 2, 3$ ; in particular, all are nonnegative numbers. This simply means that investment toward mutual assessment can be zero (i.e., self-assessment) or limitlessly greater (although increasing investment comes at greater cost), and that the withdrawal strategy against a matched opponent can vary from always withdraw ( $u_i = 0$  or  $v_i = 0$  for  $i \geq 2$ ) to always persist ( $u_i, v_i = 1, i \geq 2$ ).

We explore the model by assigning player 1 a mutant strategy and having it interact with a population of individuals all using the same strategy and in the role of player 2. The resultant fitness to a mutant  $u$ -strategist in a population of  $v$ -strategists, denoted by  $f(u, v)$ , is now readily calculated as the expected value—over all possible combinations of actual and perceived RHP—of the payoff to strategy  $u$  minus the costs of assessment. Details of this computation are given in appendix A; appendixes A–D available online.

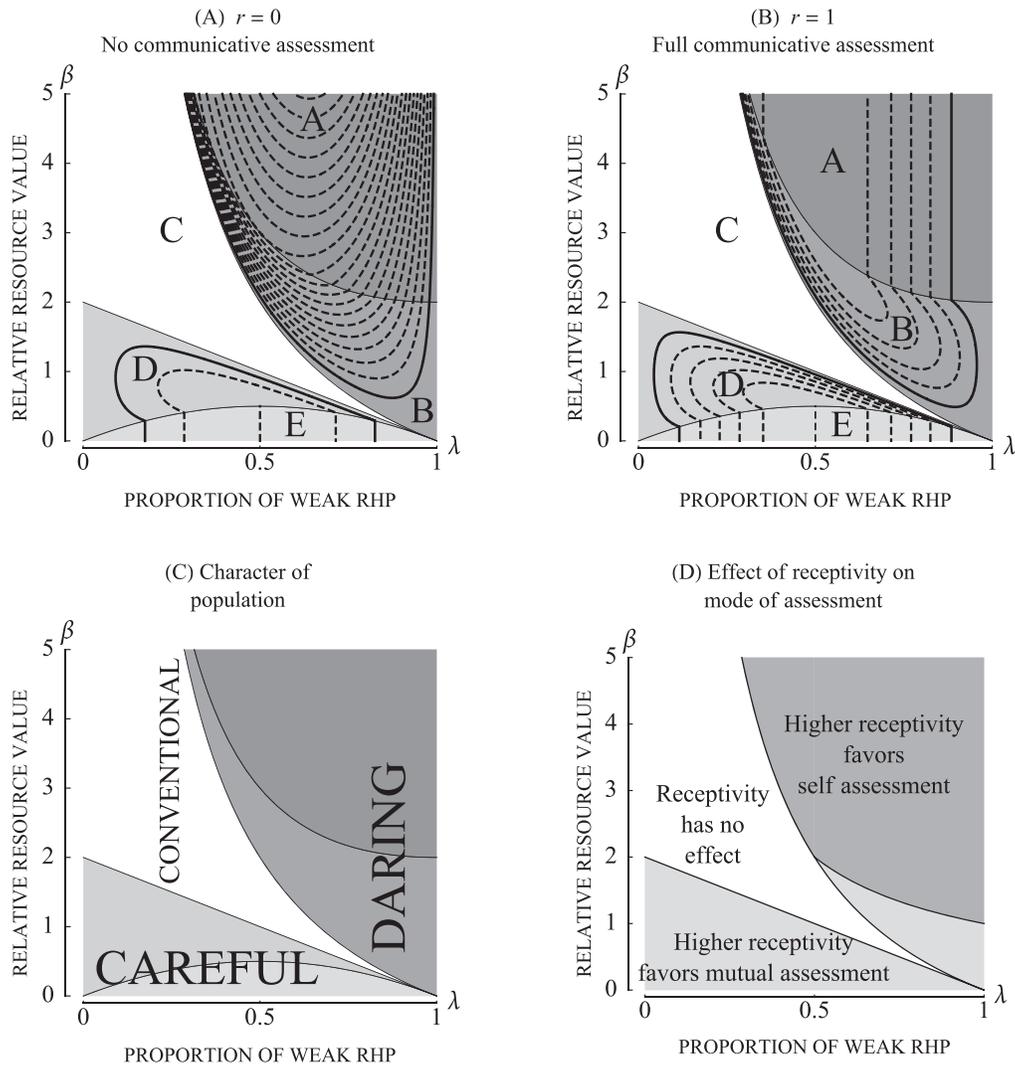
#### Evolutionarily Stable Strategy Analysis

Population strategy  $v^* = (v_1^*, v_2^*, v_3^*)$  is an evolutionarily stable strategy (ESS) in the sense of Maynard Smith (1982)

if it does not pay a potential mutant to adopt an alternative strategy  $u = (u_1, u_2, u_3)$ . The asterisk is used to distinguish the ESS  $v^*$  from an arbitrary population strategy  $v$  (see app. B). The stable population strategy is a strong ESS if it is also uniquely the best reply to itself, so that no alternative strategy can invade. The ESS is weak if there is at least one alternative strategy that does equally well against the ESS but is outperformed by the ESS in any subpopulation that adopts the alternative strategy. Intuitively, in a population at a strong ESS, any mutant strategy is immediately selected against, whereas, at a weak ESS, a strategy that is an alternative best reply can gain an initial toehold but cannot spread before being selected against. For the sake of simplicity, we assume throughout that each component of strategy is independent of the other two, so that a potential mutant strategy  $u$  differs from an ESS  $v^*$  only in one component of the vector.

Detailed ESS calculations are presented in appendix B and can be summarized as follows. First, we distinguish an ESS as type I or type II according to whether it involves pure self-assessment ( $v_1^* = 0$ ) or some form of mutual assessment ( $v_1^* > 0$ ). Second, a strategy comprises not only assessment behavior but also the conditions of withdrawal in contests. While an individual’s withdrawal behavior is fixed if it perceives an RHP asymmetry, there is the potential for different withdrawal behaviors when no asymmetry is perceived. Because a contest can be perceived as matched in two different ways (WW or SS) and there are three kinds of behavior for each (pure hawk, pure dove, or mixed), either type of ESS could have up to  $3 \times 3 = 9$  subtypes. However, only five of the possible type I subtypes and only six of the type II subtypes can arise (table 2, figs. 3, B5, app. B; figs. B1–B5 available online).

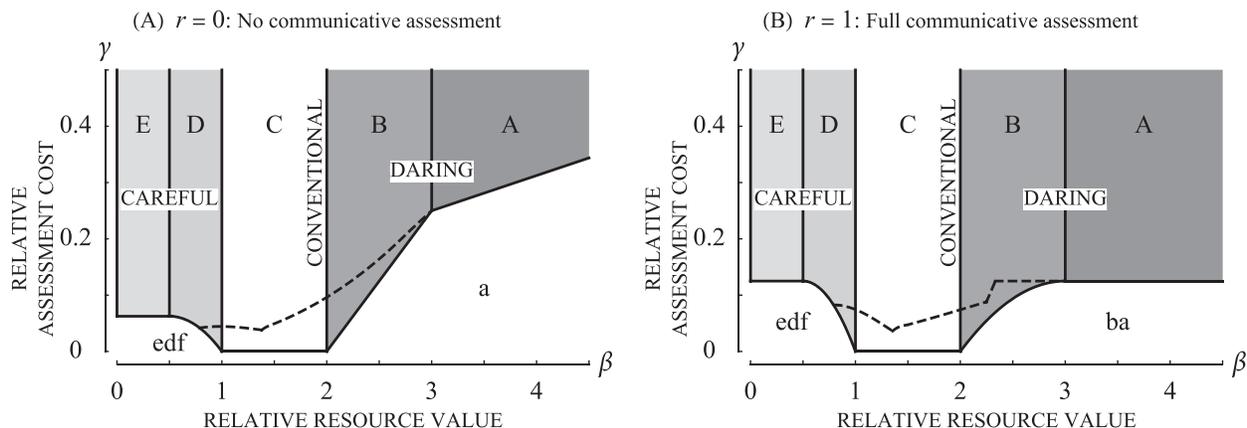
The different subtypes for the withdrawal component of strategy can be grouped to characterize one of five different types of population. First, there are “daring” populations in which strong individuals fight unconditionally



**Figure 3:** Five regions of the  $\lambda$ - $\beta$  plane corresponding to five different types of pure self-assessment evolutionarily stable strategy, where  $\lambda$  denotes proportion of weak resource holding potential (RHP) and  $\beta$  denotes relative resource value (table 1). The regions are distinguished by their shading and are defined in table 2. Superimposed are contour maps of the surface of the critical relative cost threshold ( $\gamma_c$ ) for  $r = 0$  (A) and  $r = 1$  (B), where  $r$  denotes receptivity. Increments of  $1/48$  are used, with the first contour given as a solid line. C, Regions A and B correspond to a daring population, region C to a conventional one, and regions D and E to a careful one. D, How the effect of receptivity on mode of assessment varies with RHP proportion and relative resource value.

and weak individuals are willing to fight against opponents they assess as weak (table 2a, 2b). Second, there are “careful” populations in which weak individuals always withdraw from an interaction and strong individuals persist unconditionally only against perceived weak opponents (table 2d, 2e). Third, there are populations that have adopted a “convention” for deciding contests (table 2c). In these cases, strong individuals always persist, and weak individuals always withdraw. There can also be “mixed” populations, intermediate between careful and daring (table 2f). In such populations, neither strong nor weak in-

dividuals persist or withdraw unconditionally. Finally, there can be an “anomalous” population in which weak individuals may fight among each other but strong individuals persist only against weak individuals (table 2g). Such populations occur only when strong individuals are sufficiently rare and relative resource value is sufficiently low, allowing weak individuals to be more daring and strong individuals to be more careful. Note that in all cases, a weak individual will always withdraw if it assesses its opponent to be strong, and a strong individual will always be willing to persist if it assesses its opponent as weak.



**Figure 4:** Effect of receptivity ( $r$ ) on a self-assessment evolutionarily stable strategy (ESS) when variance in resource holding potential is maximal ( $\lambda = 1/2$ ) for  $r = 0$  (A) and  $r = 1$  (B). Here, relative resource value ( $\beta$ ; table 1) increases horizontally, while relative assessment cost ( $\gamma$ ; table 1) increases vertically. There cannot exist a type I ESS in the unshaded regions below the solid boundary curve, that is, where  $\gamma$  is less than the critical threshold  $\gamma_c$ ; but there is always a type I ESS where  $\gamma$  exceeds  $\gamma_c$ , that is, in a shaded region or region C. A type II ESS can exist either below the solid curve or (above it and) below the dashed curve (corresponding to the shaded regions of fig. 5). The solid curve shows that the critical relative assessment cost,  $\gamma_c$ , varies with relative resource value  $\beta$  across different regions that represent different withdrawal strategies. Regions E and D characterize a careful population, in which only strong individuals are willing to persist. All individuals are willing to persist in regions A and B, which characterize a daring population. Region C typifies a conventional strategy in which strong individuals always persist and weak individuals always withdraw. Lowercase letters refer to corresponding withdrawal strategies that are part of the mutual assessment strategy (see fig. 5). For further details, see appendix C, available online.

To aid in analysis, we define two dimensionless parameters:

$$\beta = \frac{V}{C}, \quad \gamma = \frac{\Gamma}{\alpha C}. \quad (8)$$

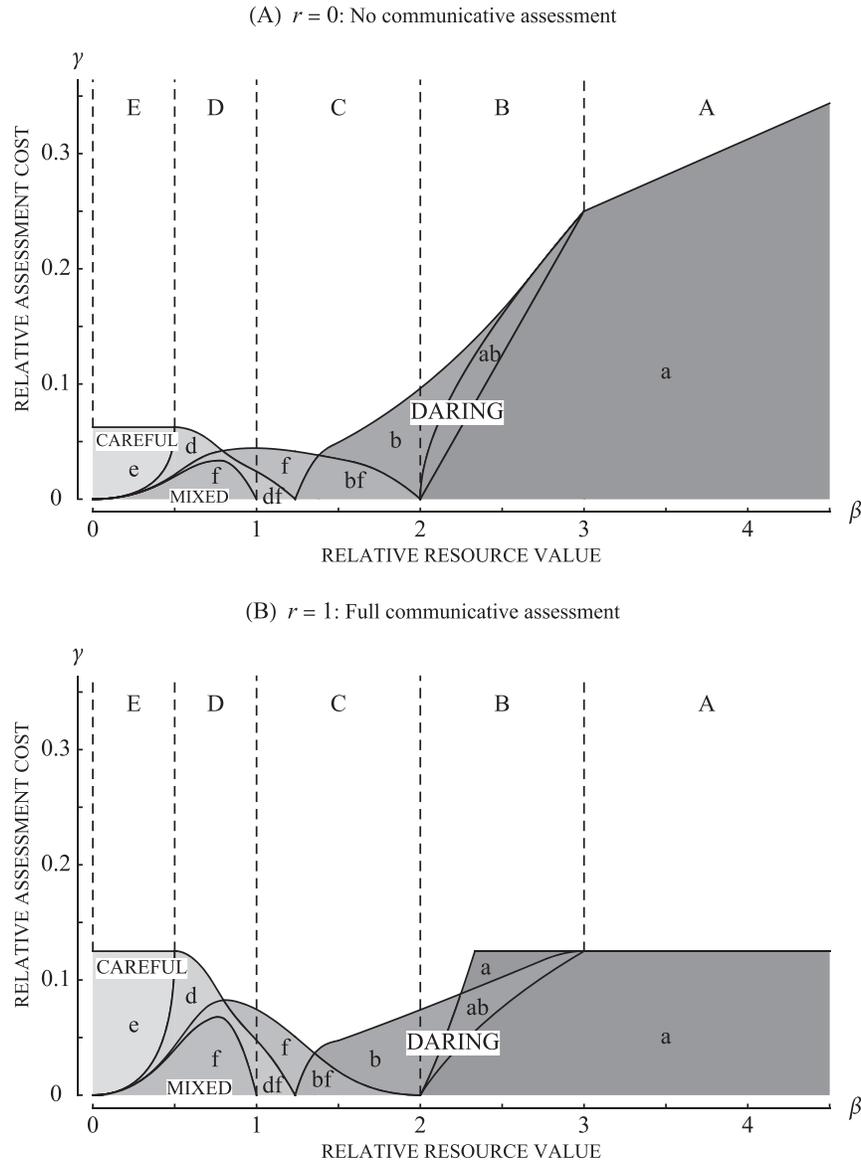
The first parameter,  $\beta$ , which we call relative resource value, measures the value of the resource relative to the cost of escalation. The second parameter,  $\gamma$ , which we call relative assessment cost, measures the effective assessment cost, which is the cost of improving assessment accuracy ( $\Gamma/\alpha$ ; see app. A), relative to the cost of escalation,  $C$ . Note that relative assessment costs may be small—and hence assessment cost-effective relative to escalation—either because information on opponents can be gained at low cost (low  $\Gamma$ ) or because it has high efficacy rates (high  $\alpha$ ). Alternatively, relative assessment costs may improve the effectiveness of gathering information if fighting is sufficiently costly (high  $C$ ). Smaller values of  $\gamma$  reflect more efficient information gain. Note that  $\beta$  and  $\gamma$  are also defined in table 1 for ease of reference.

#### *Trade-Off between Mutual Assessment and Escalation*

An individual that is behaving optimally can afford to bear only a certain amount of costs. It is in the allocation of these costs between mutual assessment and the capacity to escalate that determines whether self- or mutual assessment is the ESS for a given set of ecological conditions

defined by our parameters (table 1). Specifically, the assessment strategy used at the ESS is determined by the relative assessment cost  $\gamma$ . Mutual assessment is evolutionarily stable when the effective cost of mutual assessment is low relative to the cost of escalation, and self-assessment is evolutionarily stable when the effective cost of mutual assessment is relatively high. There is thus a critical threshold that the relative assessment cost ( $\gamma$ ) must exceed for there to be self-assessment at the ESS; otherwise, there is mutual assessment at the ESS. This critical threshold is a function of the distribution of RHP in the population ( $\lambda$ ), relative resource value ( $\beta$ ), and the degree of receptivity for communicative assessment ( $r$ ). We refer to the critical threshold as  $\gamma_c$ . For a given  $r$ , how  $\gamma_c$  varies with the RHP distribution ( $\lambda$ ) and relative resource value ( $\beta$ ) is most easily conceived as variation in the height of a surface: the lower the height, the more that self-assessment is favored. A contour map of this surface appears in figure 3A, 3B, and a cross section is in figures 4 and 5.

The shape of the surface for the critical threshold depends on the withdrawal strategy associated with the self-assessment ESS (app. B). As there are five possible withdrawal strategies that could be adopted for a self-assessment ESS, there are five regions of parameter space that define the relationship between the critical threshold, distribution of RHP, and relative resource value. These regions are independent of receptivity,  $r$ , and are presented visually in figure 3. Regions A and B reflect a daring population, while regions



**Figure 5:** Effect of receptivity ( $r$ ) on a mutual-assessment evolutionarily stable strategy (ESS) when variance in resource holding potential is maximal ( $\lambda = 1/2$ ) for  $r = 0$  (A) and  $r = 1$  (B). There can be a type II ESS only in a shaded region. Darker shading indicates a more daring population, lighter shading a more careful one. Unshaded regions correspond to the shaded regions above the dashed curve in figure 4. Regions d and e represent a careful population, in which only strong individuals are willing to persist. In regions a and b, all individuals are willing to persist, which characterizes a daring population. Also possible at the ESS are populations that adopt a mixed withdrawal strategy, classed as f. Uppercase letters refer to corresponding withdrawal strategies that are part of the self-assessment strategy (see fig. 4). For further details, see appendix C, available online.

D and E characterize a careful population. Region C represents a population that has adopted a conventional contest strategy. Note that these regions reflect the withdrawal strategy at the self-assessment ESS and that this does not necessarily imply that a population using mutual assessment has adopted the same withdrawal strategy. The importance of these regions in this context is that they define the manner

in which the critical relative cost threshold for the assessment strategy varies with RHP distribution and relative resource value.

Regardless of receptivity, the greater the height of the surface contoured in figure 3, the more that mutual assessment is favored. Thus, in general, mutual assessment is least favored in the conventional region (C) and is more

favored in the daring region (A or B) than in the careful region (D or E). That is, mutual assessment must be more cost-effective (i.e., have lower effective cost  $\Gamma/\alpha$  in relation to the costs of escalation  $C$ ) to occur in the careful region than to occur in the daring region.

The contour maps in figure 3 and the boundary line in figure 4 also indicate that there is a complex relationship between  $\gamma_c$  and  $\beta$ . With reference to figure 4, it is apparent that the critical threshold can vary positively, negatively, or not at all as  $\beta$  increases. Furthermore, the slope of  $\gamma_c$  against  $\beta$ , if positive, is steeper when receptivity ( $r$ ) is low than when receptivity is high. The implications of this relationship are that variation between self- and mutual assessment strategies in nature may depend on interactions between ecological conditions. For instance, a shift from mutual to self-assessment occurs with increasing resource value if fight costs are high (e.g., in the careful region), but if fight costs are low (e.g., in the daring region), then the same shift will occur with decreasing resource value. The case of the maximally varying population with respect to RHP ( $\lambda = 1/2$ ) corresponds to figures 4 and 5 and is analyzed in detail in appendix C. The patterns regarding variation in the critical threshold are qualitatively identical across variation in  $\lambda$ . Appreciation of the patterns in figures 4 and 5 provides an understanding of how the assessment-strategy threshold can vary with relative resource value ( $\beta$ ), which is a readily observable quantity in nature. Additionally, noting how the threshold varies with receptivity illustrates how information transfer affects adaptive strategies. Finally, a pictorial representation of the threshold aids in comprehending the trade-off between information and persistence ( $\gamma$ ).

#### *The Effect of Communication*

The potential for mutual assessment to be communicative ( $r > 0$ ) has contrasting effects on the critical relative assessment cost,  $\gamma_c$  (figs. 3–5). Specifically, there are regions in which increasing  $r$  increases the threshold  $\gamma_c$ , implying that individuals can sustain greater relative costs of mutual assessment when they are more receptive to opponent behavior. Additionally, there are regions in which increasing  $r$  decreases  $\gamma_c$ , implying that mutual assessment must be sufficiently cost-effective when individuals are receptive to one another for it to be adopted. These effects of communication can be appreciated by noticing how the critical relative cost threshold shifts as  $r$  increases from 0 to 1 in figures 3–5. In the careful region and the lowest portion of the daring region, the threshold ( $\gamma_c$ ) increases as  $r$  increases, implying that higher receptivity favors mutual assessment (fig. 3D). By contrast, in most of the daring region,  $\gamma_c$  decreases as  $r$  increases. Thus, higher receptivity favors self-assessment in the daring region.

The effect of communicative assessment depends on RHP, owing to a difference in payoffs between strong and weak individuals when opponents make strategic errors. Specifically, strong individuals benefit when opponents make fewer strategic errors and weak individuals benefit when their opponents make more strategic errors (app. A, table A2; tables A1–A3 available online). The relative benefits of opponent error to weak individuals versus the costs to strong individuals then determine whether increasing receptivity  $r$  has a positive or a negative effect on the critical threshold  $\gamma_c$ . The benefits of opponent error to weak individuals are greater in the daring region, such that increasing  $r$  has a negative effect on  $\gamma_c$  (figs. 3–5). In such cases, weak individuals can benefit more from hiding information on their own RHP and investing toward mutual assessment only when it is very cost-effective in relation to the costs of escalation. On the other hand, the costs of opponent error to strong individuals are greater within the careful region, as contests are extremely costly; hence  $r$  has a positive effect on  $\gamma_c$ . This is equivalent to individuals sharing the costs of mutual assessment through communication, allowing them to invest in mutual assessment even when it is relatively costly in comparison to the cost of escalation.

#### *Strategic Error Is Necessary for Assessment Strategy to Vary*

Notably, the critical relative cost threshold is zero within the conventional region (region C; figs. 3, 4), implying that here a self-assessment ESS exists regardless of how cost-effective acquiring specific information is in relation to fight costs. At this ESS, weak individuals completely avoid fighting and strong individuals always fight. In those circumstances, as shown in appendix B, there cannot be a type II (mutual assessment) ESS: investment toward mutual assessment provides no benefits, because the optimal withdrawal strategy has no risk of strategic error. This risk is absent because the population has adopted a pure withdrawal strategy that depends on RHP, such that weak individuals always withdraw and strong individuals always persist. Thus, weak individuals will never escalate contests that cannot be won, and strong individuals will never withdraw when victory is guaranteed. Put simply, it is impossible for mutual assessment to exist as an ESS when using a withdrawal strategy that contains no element of strategic error. However, withdrawal strategies that incorporate strategic error allow for the existence of a mutual-assessment ESS. This can be clearly seen from the potential for both a self- and a mutual-assessment ESS in particular regions (see figs. 4, 5, app. D). In these cases, the mutual-assessment ESSs are those that involve withdrawal strategies associated

with greater strategic error (i.e., more aggression in weak individuals and less aggression in strong individuals).

Therefore, strategic error is necessary for there to be variation in assessment strategy. We further explore the relationship between withdrawal and assessment strategies in appendix D.

### Discussion

There are two pertinent and simultaneous trade-offs relating to the assessment of opponents during contests. First, all individuals must decide between reducing the risks of strategic error by investing in mutual assessment, or increasing the capacity to persist in contests over resources. Second, weak individuals in a population must decide whether to acquire information about their opponents at the cost of revealing their own inferiority. Theoretically, the selection pressure underlying these trade-offs is the use of withdrawal strategies that, due to incomplete information, carry an inherent risk of responding inappropriately to an opponent but can otherwise increase access to limited resources.

Our model also demonstrates that a difference in payoffs between weak and strong individuals can significantly affect the benefits of mutual assessment. First, the group that benefits from gathering information can vary. When contests are costly in relation to the benefits of winning, weak individuals avoid fighting altogether, but strong individuals benefit from some aggression. Thus, mutual assessment is not beneficial to weak individuals, but it can allow strong individuals to avoid costly escalation against matched opponents. In contrast, when contests are relatively inexpensive, all individuals benefit by being aggressive, at least to some extent. However, weak individuals benefit by avoiding contests that they cannot win and hence benefit from mutual assessment. Strong individuals, on the other hand, can potentially win any fight and gathering information is not beneficial since they are willing to escalate unconditionally. Second, the difference between weak and strong fighters can establish a conflict over the quality of information that is transferred when assessment is communicative, which is discussed in detail below. Third, the need for mutual assessment can be avoided if individuals adopt a pure withdrawal strategy according to their RHP, which could imply that conventional solutions to contests can sufficiently reduce the costs associated with both mutual assessment and escalation. However, the use of such a strategy is restricted to particular ecological conditions.

### Variation between Self- and Mutual Assessment

The trade-off between gathering information and persisting in contests can explain why self-assessment is more common in nature than previous models predict. In order for an individual to maximize the resources it can gain from contests, it may need to adopt a withdrawal strategy that inherently carries the risk of making a strategic error. This risk can be lowered by adopting a mutual assessment strategy, but doing so reduces the individual's capacity to bear the costs associated with escalation. That is, although escalation is costly, animals may need to fight to access resources, and effort spent assessing opponents can reduce their ability to persist in a fight (e.g., Elwood and Arnott 2012). Thus individuals must compromise between reducing the risk of strategic error and maintaining the ability to engage in escalated interactions. It is the relative costs of these alternatives that determine whether there is self- or mutual assessment at the ESS. In other words, the foregone ability to afford involvement in escalated contests is the opportunity cost of mutual assessment. This can be observed in jumping spiders (*Phidippus clarus*) attempting to usurp a territory. Intruders appear to first assess the RHP of the resident and then switch to a self-assessment-based war of attrition if they assess themselves to have the RHP advantage (Kasumovic et al. 2010). This behavior implies that the intruder has reduced its investment toward obtaining specific information in favor of being able to persist in the war of attrition for longer. Our results thus contrast with the hypothesis that variation in assessment strategy is based purely on the trade-off between the costs and benefits of mutual assessment (Taylor and Elwood 2003; Prenter et al. 2006; Elias et al. 2008; Arnott and Elwood 2009). Specifically, although the costs and benefits of mutual assessment are important, it may be the comparison of this ratio to the costs of escalation that determines the stable assessment strategy rather than the ratio on its own.

Our model predicts that assessment strategy will vary according to the value of contested resources, the costs of escalation, and the cost-effectiveness of assessment. A comparison of assessment strategies used among fig wasp species suggests that assessment strategy varies with the value of resources in a fashion consistent with our model (Pereira and Prado 2005; Moore et al. 2008). In particular, male fig wasps often use self-assessment in fights over access to females in species for which females are extremely limited and hence valuable (Enquist and Leimar 1987; Murray 1987; Colegrave 1994). In contrast, mutual assessment appears to be used in species for which a given mating opportunity is less valuable, although the structure of the fig is also likely to be important (Pereira and Prado 2005; Moore et al. 2008). A shift from self- to mutual

assessment with decreasing resource value is expected under our model when contests are excessively costly, which conforms to fig wasp contests that carry the risk of serious injury and fatalities. Our model also predicts that assessment strategy can shift from mutual to self-assessment with an increase in escalation costs when relative resource value is high and individuals are typically aggressive against matched opponents. This prediction is supported by numerous observations of switching assessment strategies from mutual to self-assessment as the contest escalates from pure display to—presumably more costly—physical interactions (Morrell et al. 2005; Elias et al. 2008; Hsu et al. 2008; Kasumovic et al. 2010). Finally, Elias et al. (2008) showed that mutual assessment appeared to increase in contests between jumping spiders (*P. clarus*) that had previous experience with one another when compared to the first contest between the same individuals. This result is consistent with our model if experience improves the efficacy, and hence the cost-effectiveness, of mutual assessment.

#### *The Effect of Communicative Assessment*

The degree to which mutual assessment behavior provides opponents with information may have an important influence on the opportunity costs faced by an individual. Specifically, our model predicts that the transfer of specific information as a result of making an assessment (i.e., communicative assessment) can either increase or decrease the critical threshold that determines assessment strategy. The root of this differential effect comes from our assumptions regarding variation in the RHP of the population and the payoffs of interactions. Specifically, weak individuals receive greater expected payoffs if their opponents make a strategic error, while strong individuals benefit more if their opponents make correct decisions. This assumption is supported by observations of contests in convict cichlids (*Amatitlania nigrofasciata*): opponents that assess an individual to be weak are more likely to escalate than opponents that assess an individual to be strong (Leiser et al. 2004). Thus, weak individuals may suffer, and strong individuals benefit, from sharing information on their RHP with their opponent. Consequently, weaker individuals benefit from hiding information on their own RHP and may thus avoid communicative mutual assessment. On the other hand, stronger individuals benefit from sharing information on their own RHP, and thus, communicative assessment discounts their investment toward mutual assessment. In other words, if assessment is communicative, weak individuals face a trade-off between gaining information on opponents and providing opponents with information, while strong individuals do not. This information conflict can then influence the functional

significance of self- and mutual assessment strategies, assuming that assessment strategy is independent of RHP.

Communication of RHP between contesting individuals has traditionally been considered as a cooperative behavior that allows contestants to resolve a conflict without resorting to excessively costly outcomes (Clutton-Brock and Albon 1979; Crespi 1986; Rubenstein and Hack 1992; Keeley and Grant 1993; Hurd 1997). For instance, the sequential assessment model predicts that contestants should fully cooperate on sharing information on RHP asymmetries during contests and that information transfer is independent of RHP (Enquist and Leimar 1983; Enquist et al. 1990). However, our results suggest that, while individuals may share a common interest in avoiding costly outcomes, there can be conflicts of interest over the quality of information being transferred that depend on RHP and the evolution of cooperative mutual assessment will be limited by this conflict. While this is a novel interpretation of our model, it can offer explanations for some puzzling behaviors. For example, Enquist et al. (1990) observed contests in the cichlid *Nannacara anomala* and found that, although both contestants signaled their RHP to their opponents, the smaller contestant of a pair performed fewer signals that gave accurate information than the larger contestant. This hypothesis can also provide an explanation for theoretically confusing hyperaggression in smaller individuals, termed “Napoleons” (Just et al. 2007; Svensson et al. 2012): weak individuals may rapidly escalate in order to prevent their opponents from making an accurate assessment of the RHP asymmetry and force their stronger opponents into making a premature withdrawal decision.

#### *Testing the Assumptions and Predictions of the Model*

Although the model conforms to our current observations and can even resolve some points of contention, there is much that remains purely theoretical. In this respect, our model can provide testable hypotheses for further experimentation on natural cases where assessment strategy has been observed to vary (e.g., Morrell et al. 2005; Elias et al. 2008; Hsu et al. 2008; Kasumovic et al. 2010). Specifically, we predict that contestants will switch from mutual to self-assessment in high-cost contests as resource value increases or contest costs decrease. In contrast, we expect the opposite pattern in low-cost contests. A switch from mutual to self-assessment is expected whenever the cost-effectiveness of assessment decreases, regardless of resource value and contest costs.

Our assumption that weak individuals benefit from opponent error and strong individuals do not also requires experimental tests. We predict that weak individuals will attempt to limit information transfer (e.g., by performing fewer or less informative communicative behaviors). Ad-

ditionally, attempts by weak individuals to hide information will increase with the value of the contested resource and decrease with the costs of persistence. The information-conflict hypothesis also predicts that the performance of communicative assessment behaviors will be associated with the absolute RHP of a given individual, rather than the relative RHP of the contestants. Finally, contests that rely on communicative assessment (e.g., sequential assessment) may be less likely in populations where the RHP distribution is skewed toward weaker individuals.

#### *Mutual Assessment and Withdrawal Strategies*

From a theoretical standpoint, there are some notable differences between the results of our model and Crowley (2000) that could indicate the influence of obtaining specific information on withdrawal strategies. Specifically, Crowley (2000) shows that the ESS withdrawal strategy is an RHP-dependent pure strategy if individuals have access to only general information, there are only two RHP categories that are equally prevalent in the population, and strong individuals always defeat weak individuals. That is, strong individuals always play hawk, and weak individuals always play dove. However, we find that mixed strategies, in which individuals play a mixture of hawk and dove, can exist under the same conditions, although this outcome depends on relative resource value. Thus, the option to invest in obtaining specific information may result in the adoption of a mixed ESS for withdrawal strategies even if specific information is not used. Alternatively, this difference may be due to Crowley's (2000) inclusion of an escalation cost asymmetry between winners and losers, or the costs of sharing when both individuals play dove. The implication of our result is that the potential to invest in obtaining specific information (even if it is not obtained) can disrupt the use of a strategy in which individuals fight if they are above an RHP threshold and withdraw if they are below, in favor of a strategy in which withdrawal decisions are made on the basis of probability.

#### *Conclusions and Future Developments*

Our model can be expanded to explore the implications of continuous variation in RHP and the importance of RHP in determining contest outcomes (Crowley 2000; McNamara and Houston 2005). Furthermore, the costs of contests can be made to vary depending on the assessment strategy being used, the outcome of the contest, and the RHP of the contestants (e.g., the cost of a contest may increase as RHP asymmetry decreases for mutual assessment; losers may pay greater escalation costs; high-RHP individuals may inflict, and be able to receive, more costs;

Enquist et al. 1990; Crowley 2000; Leiser et al. 2004; Hsu et al. 2008). Similarly, the costs of assessment could vary with an individual's RHP or the asymmetry between contestants (Enquist and Leimar 1983; Enquist et al. 1990; Leiser et al. 2004). The implications of such alternative RHP conditions and cost structures must be explored to account for a more accurate representation of nature and to improve our understanding of assessment strategies.

In this article, we proposed that variation in assessment strategy is the result of a balance between investment toward mutual assessment and the capacity to escalate. However, there may be other important limitations on mutual assessment (Whitehouse 1997; Taylor et al. 2001; Taylor and Elwood 2003; Prenter et al. 2006; Elias et al. 2008; Arnott and Elwood 2009). First, there may be constraints to information processing—from cognitive abilities, available information, and sensory acuity—that limit mutual assessment (Taylor and Elwood 2003; Prenter et al. 2006; Elias et al. 2008; Arnott and Elwood 2009). For instance, a comparison of RHP may be too cognitively complex for many animals, and private and public information on RHP may instead be assessed independently and combined rather than directly compared (Elwood and Arnott 2012). However, there has been little work in terms of identifying the cognitive requirements for different ways of assessing RHP in contests. Second, investment toward mutual assessment may reduce an individual's RHP, or individuals may have to initially gather experience in contests before being able to make an accurate assessment (Arnott and Elwood 2009). Third, there is even the potential for no role assessment (neither self- nor mutual) to occur in contests, such as if individuals fight indiscriminately (Colegrave 1994; Cook and Bean 2006; Moore et al. 2008) or are able to adjust their withdrawal strategy according to experience (Whitehouse 1997). Finally, most contest models assume that individuals compare information on their own RHP to public information on opponent RHP. However, some animals appear to compete in a way that suggests private information is ignored and withdrawal decisions are made entirely on the basis of specific information without the comparative process of mutual assessment (Wise and Jaeger 1998; Rillich et al. 2007; Prenter et al. 2008; Arnott and Elwood 2009). Thus, future theoretical developments may need to consider the use of general and specific information independently of a comparison with private information (e.g., Elwood and Arnott 2012). Overall, it is clear that we must consider a wide range of processes that extend beyond the scope and capabilities of this model in order to understand assessment strategies.

The adaptive value of information-gathering strategies like mutual assessment depends on the costs of assessment. While our model considers the costs of gathering specific

information on opponent RHP, we assumed that individuals possess perfect private and general information without cost. However, specialized strategies that ignore information can be adaptive over strategies that gather information (Stephens 1989). Indeed, we predict this strategy here with respect to the evolution of gathering specific information. Thus, we may expect individuals to ignore private and general information of RHP in some cases. Generally, previous theoretical work has demonstrated that withdrawal strategies approach the solution to the original, symmetrical case of the hawk-dove game as the information in the system decreases (Crowley 2000; McNamara and Houston 2005). That is, all individuals play hawk if the value of the contest outweighs the costs of fighting and otherwise the same mixed strategy (Crowley 2000). Given that we are already familiar with the case of no private information in contests, future theoretical work must develop models that determine the conditions in which it is adaptive to ignore or gather information on own RHP. The results can then be compared to the case covered by our model, in which it is information on opponent RHP that is costly to gather. Empirical studies can then focus on finding support for any of these models.

Despite the broad scope concerning assessment strategies, our model supports an important role for strategic error at the heart of contest behavior through its influence on withdrawal strategies (Parker and Rubenstein 1981), aggressiveness (Just et al. 2007), and opponent assessment (Enquist and Leimar 1983; Just et al. 2007; this study). Furthermore, asymmetries in the payoffs of sharing information between contestants may lead to conflicts of information sharing, which could have important implications for any model that considers the use of public information. Together, these factors may serve as central components of future models that attempt to provide a general explanation of variation in assessment strategies.

#### Acknowledgments

We are very grateful to R. Bonduriansky, J. L. Bronstein, R. Elwood, C. McLean, D. Stuart-Fox, and two anonymous reviewers for constructive criticism of earlier versions of the manuscript. S.M.H. was funded by the Finnish Centre of Excellence in Biological Interactions, the John and Betty McCreery Travelling Scholarship, the Melbourne Abroad Travelling Scholarship, and the Drummond Award from the University of Melbourne. This work was partially supported by a grant from the Simons Foundation (274041 to M.M.-G.).

#### Literature Cited

- Arnott, G., and R. W. Elwood. 2008. Information gathering and decision making about resource value in animal contests. *Animal Behaviour* 76:529–542.
- . 2009. Assessment of fighting ability in animal contests. *Animal Behaviour* 77:991–1004.
- Bridge, A. P., R. W. Elwood, and J. T. A. Dick. 2000. Imperfect assessment and limited information preclude optimal strategies in male-male fights in the orb-weaving spider *Metellina mendei*. *Proceedings of the Royal Society B: Biological Sciences* 267:273–279.
- Briffa, M., and R. W. Elwood. 2009. Difficulties remain in distinguishing between mutual and self-assessment in animal contests. *Animal Behaviour* 77:759–762.
- Briffa, M., and L. U. Sneddon. 2007. Physiological constraints on contest behaviour. *Functional Ecology* 21:627–637.
- Clutton-Brock, T. H., and S. D. Albon. 1979. The roaring of red deer and the evolution of honest advertisement. *Behaviour* 69:145–170.
- Colegrave, N. 1994. Game theory models of competition in closed systems: asymmetries in fighting and competitive ability. *Oikos* 71:499–505.
- Cook, J. M., and D. Bean. 2006. Cryptic male dimorphism and fighting in a fig wasp. *Animal Behaviour* 71:1095–1101.
- Copeland, D. L., B. Levay, B. Sivaraman, C. Beebe-Fugloni, and R. L. Earley. 2011. Metabolic costs of fighting are driven by contest performance in male convict cichlid fish. *Animal Behaviour* 82:271–280.
- Crespi, B. J. 1986. Size assessment and alternative fighting tactics in *Elaphrothrips tuberculatus* (Insecta: Thysanoptera). *Animal Behaviour* 34:1324–1335.
- Crowley, P. H. 2000. Hawks, doves, and mixed-symmetry games. *Journal of Theoretical Biology* 204:543–563.
- Dall, S. R. X., L. Giraldeau, O. Olsson, J. M. McNamara, and D. W. Stephens. 2005. Information and its use by animals in evolutionary ecology. *Trends in Ecology and Evolution* 20:187–193.
- Dawkins, M., and T. Guilford. 1991. The corruption of honest signalling. *Animal Behaviour* 41:865–873.
- Elias, D. O., M. M. Kasumovic, D. Punzalan, M. C. B. Andrade, and A. C. Mason. 2008. Assessment during aggressive contests between male jumping spiders. *Animal Behaviour* 76:901–910.
- Elwood, R. W., and G. Arnott. 2012. Understanding how animals fight with Lloyd Morgan's canon. *Animal Behaviour* 84:1095–1102.
- Enquist, M., and O. Leimar. 1983. Evolution of fighting behaviour: decision rules and assessment of relative strength. *Journal of Theoretical Biology* 102:387–410.
- . 1987. Evolution of fighting behaviour: the effect of variation in resource value. *Journal of Theoretical Biology* 127:187–205.
- Enquist, M., O. Leimar, T. Ljungberg, Y. Mallner, and N. Segerdahl. 1990. A test of the sequential assessment game: fighting in the cichlid fish *Nannacara anomala*. *Animal Behaviour* 40:1–14.
- Giraldeau, L.-A. 1997. The ecology of information use. Pages 42–68 in J. R. Krebs and N. B. Davies, eds. *Behavioural ecology: an evolutionary approach*. 4th ed. Blackwell Science, Oxford.
- Hack, M. A. 1997a. Assessment strategies in the contests of male crickets, *Acheta domesticus* (L.). *Animal Behaviour* 53:733–747.
- . 1997b. The energetic costs of fighting in the house cricket, *Acheta domesticus* L. *Behavioral Ecology* 8:28–36.
- Haley, M. P. 1994. Resource-holding power asymmetries, the prior residence effect, and reproductive payoffs in male northern elephant seal fights. *Behavioral Ecology and Sociobiology* 34:427–434.

- Hammerstein, P., and G. A. Parker. 1982. The asymmetric war of attrition. *Journal of Theoretical Biology* 96:647–682.
- Hsu, Y., R. L. Earley, and L. L. Wolf. 2006. Modulation of aggressive behaviour by fighting experience: mechanisms and contest outcomes. *Biological Reviews* 81:33–74.
- Hsu, Y., S.-P. Lee, M.-H. Chen, S.-Y. Yang, and K.-C. Cheng. 2008. Switching assessment strategy during a contest: fighting in killifish *Kryptolebias marmoratus*. *Animal Behaviour* 75:1641–1649.
- Hurd, P. L. 1997. Is signaling of fighting ability costlier for weaker individuals? *Journal of Theoretical Biology* 184:83–88.
- Jakobsson, S., O. Brick, and C. Kullberg. 1995. Escalated fighting behaviour incurs increased predation risk. *Animal Behaviour* 49:235–239.
- Jennings, D. J., M. P. Gammell, R. J. H. Payne, and T. J. Hayden. 2005. An investigation of assessment games during fallow deer fights. *Ethology* 111:511–526.
- Just, W., M. R. Morris, and X. Sun. 2007. The evolution of aggressive losers. *Behavioural Processes* 74:342–350.
- Kasumovic, M. M., A. C. Mason, M. C. B. Andrade, and D. O. Elias. 2010. The relative importance of RHP and resource quality in contests with ownership asymmetries. *Behavioral Ecology* 22:39–45.
- Keeley, E. R., and J. W. A. Grant. 1993. Visual information, resource value, and sequential assessment in convict cichlid (*Cichlasoma nigrofasciatum*) contests. *Behavioral Ecology* 4:345–349.
- Kelly, C. D. 2006. Fighting for harems: assessment strategies during male-male contests in the sexually dimorphic Wellington tree weta. *Animal Behaviour* 72:727–736.
- Kokko, H. 2013. Dyadic contests: modelling fights between two individuals. Pages 5–32 in I. C. W. Hardy and M. Briffa, eds. *Animal contests*. Cambridge University Press, Cambridge.
- Leiser, J. K., J. L. Gagliardi, and M. Itzkowitz. 2004. Does size matter? assessment and fighting in small and large size-matched pairs of adult male convict cichlids. *Journal of Fish Biology* 64:1339–1350.
- Maynard Smith, J. 1974. The theory of games and the evolution of animal conflicts. *Journal of Theoretical Biology* 47:209–221.
- . 1982. *Evolution and the theory of games*. Cambridge University Press, Cambridge.
- . 2000. The concept of information in biology. *Philosophy of Science* 67:177–194.
- Maynard Smith, J., and G. A. Parker. 1976. The logic of asymmetric contests. *Animal Behaviour* 24:159–175.
- McNamara, J. M., and A. I. Houston. 2005. If animals know their own fighting ability, the evolutionarily stable level of fighting is reduced. *Journal of Theoretical Biology* 232:1–6.
- Mesterton-Gibbons, M., J. H. Marden, and L. A. Dugatkin. 1996. On wars of attrition without assessment. *Journal of Theoretical Biology* 181:65–83.
- Moore, J. C., D. J. Obbard, C. Reuter, S. A. West, and J. M. Cook. 2008. Fighting strategies in two species of fig wasp. *Animal Behaviour* 76:315–322.
- Morrell, L. J., P. R. Y. Backwell, and N. B. Metcalfe. 2005. Fighting in fiddler crabs *Uca mjoebergi*: what determines duration? *Animal Behaviour* 70:653–662.
- Murray, M. G. 1987. The closed environment of the fig receptacle and its influence on male conflict in the Old World fig wasp *Philotrypes pilosa*. *Animal Behaviour* 35:488–506.
- . 1990. Comparative morphology and mate competition of flightless male fig wasps. *Animal Behaviour* 39:434–443.
- Parker, G. A. 1974. Assessment strategy and the evolution of fighting behaviour. *Journal of Theoretical Biology* 47:223–243.
- Parker, G. A., and D. I. Rubenstein. 1981. Role assessment, reserve strategy, and acquisition of information in asymmetric animal conflicts. *Animal Behaviour* 29:221–240.
- Payne, R. J. H. 1998. Gradually escalating fights and displays: the cumulative assessment model. *Animal Behaviour* 56:651–662.
- Payne, R. J. H., and M. Pagel. 1996. When is false modesty a false economy? an optimality model of escalating signals. *Proceedings of the Royal Society B: Biological Sciences* 263:1545–1550.
- . 1997. Why do animals repeat displays? *Animal Behaviour* 54:109–119.
- Pereira, R. A. S., and A. P. D. Prado. 2005. Recognition of competitive asymmetries reduces the severity of fighting in male *Idarnes* fig wasps. *Animal Behaviour* 70:249–256.
- Prenter, J., R. W. Elwood, and P. W. Taylor. 2006. Self-assessment by males during energetically costly contests over precopula females in amphipods. *Animal Behaviour* 72:861–868.
- Prenter, J., P. W. Taylor, and R. W. Elwood. 2008. Large body size for winning and large swords for winning quickly in swordtail males, *Xiphophorus helleri*. *Animal Behaviour* 75:1981–1987.
- Rillich, J., K. Schildberger, and P. A. Stevenson. 2007. Assessment strategy of fighting crickets revealed by manipulating information exchange. *Animal Behaviour* 74:823–836.
- Rubenstein, D. I., and M. A. Hack. 1992. Horse signals? the sounds and scents of fury. *Evolutionary Ecology* 6:254–260.
- Smallegange, I., M. Sabelis, and J. Vandermeer. 2007. Assessment games in shore crab fights. *Journal of Experimental Marine Biology and Ecology* 351:255–266.
- Stephens, D. W. 1989. Variance and the value of information. *American Naturalist* 134:128–140.
- Stuart-Fox, D. 2006. Testing game theory models: fighting ability and decision rules in chameleon contests. *Proceedings of the Royal Society B: Biological Sciences* 273:1555–1561.
- Stuart-Fox, D., D. Firth, A. Moussalli, and M. J. Whiting. 2006. Multiple signals in chameleon contests: designing and analysing animal contests as a tournament. *Animal Behaviour* 71:1263–1271.
- Svensson, P. A., T. K. Lehtonen, and B. B. M. Wong. 2012. A high aggression strategy for smaller males. *PLoS ONE* 7:e43121.
- Taylor, P. W., and R. W. Elwood. 2003. The mismeasure of animal contests. *Animal Behaviour* 65:1195–1202.
- Taylor, P. W., O. Hasson, and D. L. Clark. 2001. Initiation and resolution of jumping spider contests: roles for size, proximity, and early detection of rivals. *Behavioral Ecology and Sociobiology* 50:403–413.
- Wagner, R. H., and É. Danchin. 2010. A taxonomy of biological information. *Oikos* 119:203–209.
- West, S., M. G. Murray, C. A. Machado, A. S. Griffin, and E. A. Herre. 2001. Testing Hamilton's rule with competition between relatives. *Nature* 409:510–513.
- Whitehouse, M. E. A. 1997. Experience influences male-male contests in the spider *Argyrodes antipodiana* (Theridiidae: Araneae). *Animal Behaviour* 53:913–923.
- Wise, S. E., and R. G. Jaeger. 1998. The influence of tail autotomy on agonistic behaviour in a territorial salamander. *Animal Behaviour* 55:1707–1716.

# Appendix A from M. Mesterton-Gibbons and S. M. Heap, “Variation between Self- and Mutual Assessment in Animal Contests” (Am. Nat., vol. 183, no. 2, p. 199)

## Calculation of Fitness

Let  $q_{ij}(X, Y)$  denote the probability that the perceptual combination is  $j$  when the actual RHP combination is  $i$  and the total effective effort expended on mutual assessment is  $X$  for player 1 but  $Y$  for player 2, according to equations (3) and (4). Then the resultant  $4 \times 4$  matrix is  $\mathbf{Q}(X, Y)$ , defined by table A1.

Let  $a_{ij}$  be the payoff to player 1 from an interaction with player 2 when the RHP combination is  $i$  and perceived as  $j$ . Then the  $4 \times 4$  matrix  $\mathbf{A}$  is calculated as follows. First, suppose that both individuals are weak and that each correctly perceives the other as weak. Then player 1 escalates with probability  $u_2$  and is nonaggressive with probability  $1 - u_2$ . Because the corresponding probabilities for player 2 are  $v_2$  and  $1 - v_2$ , the payoff to player 1 in this instance is  $a_{11} = \{(1/2)V - C\} \cdot u_2 v_2 + V \cdot u_2(1 - v_2) + 0 \cdot (1 - u_2)v_2 + (1/2)V \cdot (1 - u_2)(1 - v_2) = (1/2)V\{1 - v_2 + u_2\} - Cu_2 v_2$ . If instead player 1 correctly perceives player 2 as weak but player 2 misperceives player 1 as strong, then because a strong animal is guaranteed to win a fight, player 2 does not escalate; and so the payoff to player 1 is  $V$  if it escalates and  $(1/2)V$  if it is nonaggressive; that is,  $a_{12} = V \cdot u_2 + (1/2)V \cdot (1 - u_2) = (1/2)V\{1 + u_2\}$ . Similarly,  $a_{13} = 0 \cdot v_2 + (1/2)V(1 - v_2)$  and  $a_{14} = (1/2)V$ .

Next suppose that player 1 is weak but player 2 is strong, so that if there is a fight, then player 2 is guaranteed to win. However, there will be no fight if both perceive correctly: player 2 will obtain the resource without cost to either contestant, and so  $a_{21} = 0$ . If, on the other hand, both perceive incorrectly, that is, player 1 perceives player 2 as weak while player 2 perceives player 1 as strong, then player 1 will escalate with probability  $u_2$  while player 2 escalates with probability  $v_3$ , and so  $a_{24} = -C \cdot u_2 v_3 + V \cdot u_2(1 - v_3) + 0 \cdot (1 - u_2)v_3 + (1/2)V \cdot (1 - u_2)(1 - v_3) = (1/2)V(1 + u_2)(1 - v_3) - Cu_2 v_3$ . Similarly,  $a_{22} = 0 \cdot v_3 + (1/2)V \cdot (1 - v_3)$  and  $a_{23} = -C \cdot u_2 + 0 \cdot (1 - u_2)$ .

Next suppose that player 1 is strong but player 2 is weak, so that player 1 is guaranteed to win any fight. There will be no fight if both perceive correctly: player 1 will obtain the resource without cost, and so  $a_{31} = V$ . If, on the other hand, both perceive incorrectly, that is, player 1 perceives player 2 as strong while player 2 perceives player 1 as weak, then player 1 will escalate with probability  $u_3$  while player 2 escalates with probability  $v_2$ , and so  $a_{34} = \{V - C\} \cdot u_3 v_2 + V \cdot u_3(1 - v_2) + 0 \cdot (1 - u_3)v_2 + (1/2)V \cdot (1 - u_3)(1 - v_2) = (1/2)V\{1 - v_2 + (1 + v_2)u_3\} - Cu_3 v_2$ . Similarly,  $a_{32} = \{V - C\} \cdot v_2 + V \cdot (1 - v_2) = V - Cv_2$  and  $a_{33} = V \cdot u_3 + (1/2)V(1 - u_3) = (1/2)V(1 + u_3)$ .

Finally, suppose that both individuals are strong, and that each correctly perceives the other as strong. Then player 1 escalates with probability  $u_3$  while player 2 escalates with probability  $v_3$ , and so  $a_{41} = \{(1/2)V - C\} \cdot u_3 v_3 + V \cdot u_3(1 - v_3) + 0 \cdot (1 - u_3)v_3 + (1/2)V \cdot (1 - u_3)(1 - v_3) = (1/2)V\{1 - v_3 + u_3\} - Cu_3 v_3$ . Similarly,  $a_{42} = \{(1/2)V - C\}u_3 + 0 \cdot (1 - u_3)$ ,  $a_{43} = \{(1/2)V - C\}v_3 + V \cdot (1 - v_3)$  and—because a fight is inevitable if both misperceive— $a_{44} = (1/2)V - C$ . We thus obtain the  $4 \times 4$  contest payoff matrix  $\mathbf{A}$  presented in table A2. These payoffs exclude the cost of assessment.

However, the total fitness to a  $u$ -strategist in a population of  $v$ -strategists includes the cost of assessment. We obtain

$$f(u, v) = \sum_{i=1}^4 \sum_{j=1}^4 a_{ij} p_i q_{ij}(X, Y) - \Gamma u_1, \quad (\text{A1})$$

where  $p_i$  is defined by equation (2),  $X = u_1 + r v_1$  and  $Y = r u_1 + v_1$  by equations (3) and (4),  $q_{ij}(X, Y)$  by table A1, and  $a_{ij}$  by table A2. After simplification, (A1) reduces to

$$f(u, v) = \left\{ \frac{1}{2} \beta B_1(u, v) - B_2(u, v) - e^{-(1+r)(u_1+v_1)\alpha} B_3(u, v) - \alpha \gamma u_1 \right\} C, \quad (\text{A2})$$

with  $B_1$ ,  $B_2$ , and  $B_3$  newly defined by table A3, whereas  $\lambda = V/C$  and  $\gamma = (\Gamma/\alpha)/C$  were earlier defined by equation (8). As noted in the main body of the article,  $\gamma$  measures the effective assessment cost relative to the cost of escalation. (To see why  $\Gamma/\alpha$  is the effective cost of assessment, note from equation [6] that  $E = 1/\alpha$  is the amount of effort required to reduce relative perception error to  $e^{-1}$  or 37%; and because the cost per unit effort is  $\Gamma$ , the cost of reducing relative error to 37% is  $\Gamma \cdot 1/\alpha$ .)

**Table A1.** The matrix  $\mathbf{Q}(X, Y)$ , which yields the probabilities associated with all 16 possible combinations of actual and perceptual state for an interaction between two individuals

State $i$	RHP		Probability (CC) $q_{i1}$	Probability (CI) $q_{i2}$	Probability (IC) $q_{i3}$	Probability (II) $q_{i4}$
	$P_1$	$P_2$				
1	W	W	$\phi_w(X)\phi_w(Y)$	$\phi_w(X)\{1-\phi_w(Y)\}$	$\{1-\phi_w(X)\}\phi_w(Y)$	$\{1-\phi_w(X)\}\{1-\phi_w(Y)\}$
2	W	S	$\phi_s(X)\phi_w(Y)$	$\phi_s(X)\{1-\phi_w(Y)\}$	$\{1-\phi_s(X)\}\phi_w(Y)$	$\{1-\phi_s(X)\}\{1-\phi_w(Y)\}$
3	S	W	$\phi_w(X)\phi_s(Y)$	$\phi_w(X)\{1-\phi_s(Y)\}$	$\{1-\phi_w(X)\}\phi_s(Y)$	$\{1-\phi_w(X)\}\{1-\phi_s(Y)\}$
4	S	S	$\phi_s(X)\phi_s(Y)$	$\phi_s(X)\{1-\phi_s(Y)\}$	$\{1-\phi_s(X)\}\phi_s(Y)$	$\{1-\phi_s(X)\}\{1-\phi_s(Y)\}$

Note:  $P_k$  is used as a shorthand for player  $k$ .

**Table A2.** The matrix  $\mathbf{A}$ , which yields the payoffs associated with all 16 possible combinations of actual and perceptual state for an interaction between two individuals

State $i$	RHP		CC $j = 1$	CI $j = 2$	IC $j = 3$	II $j = 4$
	$P_1$	$P_2$				
1	W	W	$(1/2)V\{1 - v_2 + u_2\} - Cu_2v_2$	$(1/2)V\{1 + u_2\}$	$(1/2)V(1 - v_2)$	$(1/2)V$
2	W	S	0	$(1/2)V(1 - v_3)$	$-Cu_2$	$(1/2)V(1 + u_2)(1 - v_3) - Cu_2v_3$
3	S	W	$V$	$V - Cv_2$	$(1/2)V(1 + u_3)$	$(1/2)V\{1 - v_2 + (1 + v_2)u_3\} - Cu_3v_2$
4	S	S	$(1/2)V\{1 - v_3 + u_3\} - Cu_3v_3$	$\{(1/2)V - C\}u_3$	$V\{1 - (1/2)v_3\} - Cv_3$	$(1/2)V - C$

Note:  $P_k$  is used as a shorthand for player  $k$ . It is readily verified that  $a_{i2} \geq a_{i1}$  and  $a_{i4} \geq a_{i3}$  for  $i \leq 2$ , whereas both inequalities are reversed for  $i \geq 3$ . Thus, weak individuals benefit when their opponents make more strategic errors, whereas strong individuals benefit when their opponents make fewer strategic errors.

**Table A3.** Definition of terms used to yield an explicit expression for fitness

$k$	$B_k(u, v)$
1	$1 + (1 - \lambda)^2(u_3 - v_3) + \lambda^2(u_2 - v_2) + (1 - \lambda)\lambda^2\{e^{-(r u_1 + v_1)\alpha}v_2 - e^{-(u_1 + r v_1)\alpha}u_2\}$ $+ (1 - \lambda)^2\lambda^2 e^{-(1+r)(u_1+v_1)\alpha}\{u_2(1 - v_3) - v_2(1 - u_3)\}$
2	$(1 - \lambda)^2\{u_3 + \lambda(1 - u_3)e^{-(u_1 + r v_1)\alpha}\}\{v_3 + \lambda(1 - v_3)e^{-(r u_1 + v_1)\alpha}\}$ $+ \lambda^2 u_2 v_2 \{1 - (1 - \lambda)e^{-(u_1 + r v_1)\alpha}\}\{1 - (1 - \lambda)e^{-(r u_1 + v_1)\alpha}\}$
3	$\lambda^2(1 - \lambda)(v_2\{e^{(u_1 + r v_1)\alpha} - (1 - \lambda)(1 - u_3)\} + u_2\{e^{(r u_1 + v_1)\alpha} - (1 - \lambda)(1 - v_3)\})$

## Appendix B from M. Mesterton-Gibbons and S. M. Heap, “Variation between Self- and Mutual Assessment in Animal Contests” (Am. Nat., vol. 183, no. 2, p. 199)

### Evolutionarily Stable Strategy (ESS) Calculations

Here we use  $u = (u_1, u_2, u_3)$  to denote the strategy of a focal individual who is a potential mutant,  $v = (v_1, v_2, v_3)$  to denote a population strategy, and  $v^* = (v_1^*, v_2^*, v_3^*)$  to denote a population strategy that is a candidate for ESS. As noted in the main body of the article,  $v^*$  is actually an ESS if it does not pay a potential mutant to adopt any alternative strategy;  $v^*$  is a strong ESS if it is also uniquely the best reply to itself, that is, if  $f(v^*, v^*) > f(u, v^*)$  for all  $u \neq v^*$  and a weak ESS if there is at least one alternative best reply  $u$ , but  $v^*$  is a better reply than  $u$  to all such  $u$ , that is,  $f(v^*, u) > f(u, u)$  for all  $u$  satisfying  $f(u, v^*) = f(v^*, v^*)$ . Also as noted already, we assume that each component of strategy is independent of the other two, and so a potential mutant strategy  $u$  differs from an ESS  $v^*$  only in one component of the vector.

A necessary condition for  $0 < v_i^* < 1$  at such an ESS is

$$\left. \frac{\partial f}{\partial u_i} \right|_{u=v=v^*} = 0. \quad (\text{B1})$$

Correspondingly, a necessary condition for  $v_i^* = 0$  at such an ESS is

$$\left. \frac{\partial f}{\partial u_i} \right|_{u=v=v^*} \leq 0, \quad (\text{B2})$$

and a necessary condition for  $v_i^* = 1$  at such an ESS is

$$\left. \frac{\partial f}{\partial u_i} \right|_{u=v=v^*} \geq 0. \quad (\text{B3})$$

Because  $f(u, v)$  is linear in both  $u_2$  and  $u_3$ , there must be alternative best replies when equation (B1) is satisfied with  $0 < v_i^* < 1$  for  $i = 2$  or  $i = 3$ , and so  $v^*$  can be a strong ESS only if it is a boundary ESS: either  $v_2^* = 0$  or  $v_2^* = 1$ , and either  $v_3^* = 0$  or  $v_3^* = 1$ . Such a  $v^*$  is indeed a strong ESS if equation (B2) or (B3) is satisfied with strict inequality for both  $i = 2$  and  $i = 3$ , and either equation (B2) for  $i = 1$  is satisfied with strict inequality (for  $v_1^* = 0$ ), or  $\partial^2 f / \partial u_1^2 |_{u=v=v^*} < 0$  (for  $v_1^* > 0$ ). Otherwise  $v^*$  can only be a weak ESS, and it must be verified directly that  $f(v^*, u) > f(u, u)$  when  $u \neq v^*$  for alternative best replies  $u$  with  $f(u, v^*) = f(v^*, v^*)$ .

Because we may have  $v_i^* = 0$  or  $0 < v_i^* < 1$  for  $i = 1, \dots, 3$  and  $v_i^* = 1$  for  $i = 2, 3$  at the ESS, there exist up to  $3^2 \times 2 = 18$  possible forms that a candidate ESS can take. We consider each in turn, starting with  $v_i^* = 0$  in section B.1 and continuing with  $v_i^* > 0$  in section B.2.

#### B.1. Self-Assessment ESS Conditions

To obtain conditions for a (pure) self-assessment ESS, we first set  $v_1^* = 0$ . Then, from equation (A2) in conjunction with table A3, we have

$$\begin{aligned} \left. \frac{\partial f}{\partial u_1} \right|_{u=v=v^*} &= \frac{1}{2} \lambda^2 (1 - \lambda) (1 - r) \alpha v_2^* V - (1 + r) \alpha (1 - \lambda) \lambda \{ (1 - \lambda)^2 v_3^{*2} - \lambda (1 - v_2^*) \{ 1 - \lambda (1 - v_2^*) \} \\ &\quad - (1 - \lambda) v_3^* \{ 1 - 2\lambda (1 - v_2^*) \} \} C - \Gamma, \end{aligned} \quad (\text{B4})$$

$$\left. \frac{\partial f}{\partial u_2} \right|_{u=v=v^*} = \lambda^2 \left\{ \frac{1}{2} \{ (1 - v_3^*) (1 - \lambda)^2 + \lambda \} V - \{ (1 - \lambda) \{ (1 - \lambda) v_3^* + \lambda \} + \lambda^2 v_2^* \} C \right\}, \quad (\text{B5})$$

and

$$\left. \frac{\partial f}{\partial u_3} \right|_{u=v=v^*} = (1 - \lambda)^2 \left\{ \frac{1}{2} (1 + \lambda^2 v_2^*) V - \{(1 - \lambda)\{(1 - \lambda)v_3^* + \lambda\} + \lambda^2 v_2^*\} C \right\}. \quad (\text{B6})$$

We now consider each of nine possibilities for  $v_1^* = 0$  in turn.

*B.1.1. Type IA ESS with  $v_1^* = 0$ ,  $v_2^* = 1$ ,  $v_3^* = 1$*

From equations (B2) with  $i = 1$ , (B3) with  $i = 2, 3$ , and (B4)–(B6), we require the inequalities

$$(1 - \lambda)\lambda^2\{(1 - r)\beta + 2(1 + r)(1 - \lambda)\} \leq 2\gamma, \quad (\text{B7a})$$

$$\lambda\beta \geq 2(1 - \lambda + \lambda^2), \quad (\text{B7b})$$

and  $(1 + \lambda^2)\beta \geq 2(1 - \lambda + \lambda^2)$ , which equation (B7b) implies, to hold simultaneously. We will refer to such an ESS as type IA. From equation (B7b), it corresponds to the upper portion of the daring region in fig. 3A–3C. When equation (B7) holds with strict inequality,  $v^*$  is a strong ESS.

*B.1.2. Type IE ESS with  $v_1^* = 0$ ,  $v_2^* = 0$ ,  $v_3^* = 0$*

From equations (B2) with  $i = 1, 2, 3$ , (B4), (B6), and (B5), we require

$$(1 + r)(1 - \lambda)^2\lambda^2 \leq \gamma, \quad (\text{B8a})$$

$$\beta \leq 2\lambda(1 - \lambda), \quad (\text{B8b})$$

and  $(1 - \lambda + \lambda^2)\beta \leq 2\lambda(1 - \lambda)$ , which equation (B8b) implies, to hold simultaneously. We will refer to such an ESS as type IE. From equation (B8b), it corresponds to the lower portion of the careful region in figure 3A–3C. When equation (B8) holds with strict inequality,  $v^*$  is a strong ESS.

*B.1.3. Type IC ESS with  $v_1^* = 0$ ,  $v_2^* = 0$ ,  $v_3^* = 1$*

Here equation (B4) reduces to  $-\Gamma$ , which is always negative. So from equations (B2) with  $i = 2$ , (B3) with  $i = 3$ , and (B5), (B6), we require

$$2(1 - \lambda) \leq \beta \leq \frac{2(1 - \lambda)}{\lambda}. \quad (\text{B9})$$

We will refer to such an ESS as type IC; it corresponds to the unshaded region of figure 3. When equation (B9) holds with strict inequality,  $v^*$  is a strong ESS.

*B.1.4. ESS with  $v_1^* = 0$ ,  $v_2^* = 1$ ,  $v_3^* = 0$*

From equations (B2) with  $i = 1, 3$ , (B3) with  $i = 2$ , and (B4)–(B6), we require both  $(1 - \lambda + \lambda^2)\beta \geq 2\lambda$  and  $(1 + \lambda^2)\beta \leq 2\lambda$ , which is impossible, because  $1 - \lambda + \lambda^2 < 1 + \lambda^2$  for  $\lambda > 0$ . Hence this case does not arise.

*B.1.5. Type ID ESS with  $v_1^* = 0$ ,  $v_2^* = 0$ ,  $0 < v_3^* < 1$*

From equations (B2) with  $i = 1$ , (B1) with  $i = 3$ , (B4), and (B6), we require

$$(1 + r)\lambda\beta \left\{ 1 - \frac{\beta}{2(1 - \lambda)} \right\} \leq 2\gamma \quad (\text{B10})$$

and

$$v_3^* = \frac{\beta - 2\lambda(1 - \lambda)}{2(1 - \lambda)^2} \quad (\text{B11})$$

with

$$2\lambda(1 - \lambda) < \beta < 2(1 - \lambda) \quad (\text{B12})$$

(and eq. [B5] reduces to  $-(1/4)\lambda^2\beta V$ , so that equation [B2] holds with strict inequality for  $i = 2$ ). We will refer to

such an ESS as type ID; it corresponds to the upper portion of the careful region in figures 3A–3C. Even if equation (B10) holds with strict inequality,  $v^*$  is only a weak ESS because there are always alternative best replies. That is, with  $u = (v_1^*, v_2^*, u_3)$ , we have  $f(u, v^*) = f(v^*, v^*) = (1/4)\beta(2 - \beta)C$ , but  $f(v^*, u) - f(u, u) = (1 - \lambda)^4(v_3^* - u_3)^2C > 0$  for all  $u_3 \neq v_3^*$ .

*B.1.6. ESS with  $v_1^* = 0$ ,  $0 < v_2^* < 1$ ,  $0 < v_3^* < 1$*

This case cannot arise, because equation (B1) would have to hold with  $i = 2$  and  $i = 3$ , implying  $v_2^* = -1/\lambda^2 < 0$  and  $v_3^* = (1 - \lambda + \lambda^2)/(1 - \lambda)^2 > 1$ .

*B.1.7. ESS with  $v_1^* = 0$ ,  $v_2^* = 1$ ,  $0 < v_3^* < 1$*

This case cannot arise because it requires  $2(1 - \lambda)^2v_3^* = (1 + \lambda^2)\beta - 2\lambda$ , by equation (B1) with  $i = 3$  and equation (B6). Hence  $\partial f/\partial u_2|_{u=v^*} = -(1/2)\lambda^2(\lambda + \{(1 - \lambda)^2 + \lambda\}v_3^*)V = -(1/4)\lambda^2(1 + \lambda^2)\beta V < 0$  by equation (B5), and so equation (B3) fails to hold for  $i = 2$ .

*B.1.8. ESS with  $v_1^* = 0$ ,  $0 < v_2^* < 1$ ,  $v_3^* = 0$*

This case cannot arise because it requires  $2\lambda^2v_2^* = (1 - \lambda + \lambda^2)\beta - 2\lambda(1 - \lambda)$ , by equation (B1) with  $i = 2$  and equation (B5). Hence  $\partial f/\partial u_3|_{u=v^*} = (1/2)\lambda(1 - \lambda)^2(1 - \lambda + \lambda v_2^*)V = (1/4)(1 - \lambda)^2(1 - \lambda + \lambda^2)\beta V > 0$  by (B6), and so (B2) fails to hold for  $i = 3$ .

*B.1.9. Type IB ESS with  $v_1^* = 0$ ,  $0 < v_2^* < 1$ ,  $v_3^* = 1$*

From equations (B1) with  $i = 2$  and (B5), we require

$$v_2^* = \frac{\lambda\beta - 2(1 - \lambda)}{2\lambda^2} \quad (\text{B13})$$

with

$$\frac{2(1 - \lambda)}{\lambda} < \beta < \frac{2(1 - \lambda + \lambda^2)}{\lambda}, \quad (\text{B14})$$

and equation (B3) holds with  $i = 3$  because  $\partial f/\partial u_3|_{u=v^*} = (1/2)(1 - \lambda)^2(1 - \lambda + \lambda^2v_2^*)V = (1/4)(1 - \lambda)^2\lambda\beta V > 0$  by equation (B6). From equations (B2) with  $i = 1$  and (B4), we also require

$$(1 - \lambda)\lambda\{1 + (3 - 2\lambda)r\}\beta \leq 2(1 + r)(1 - \lambda)^2 + 2\lambda\gamma + r(1 - \lambda)\lambda^2\beta^2. \quad (\text{B15})$$

We will refer to such an ESS as type IB; it corresponds to the lower portion of the daring region in figure 3A–3C. Even if equation (B15) holds with strict inequality,  $v^*$  is only a weak ESS because there are always alternative best replies: with  $u = (v_1^*, u_2, v_2^*)$ , we have  $f(u, v^*) = f(v^*, v^*) = (1/4)\beta(2 - \beta\lambda^2)C$ , but  $f(v^*, u) - f(u, u) = \lambda^4(v_2^* - u_2)^2C > 0$  for all  $u_2 \neq v_2^*$ .

Of the nine cases so far considered, we have now excluded four possibilities, namely,  $v_1^* = 0$ ,  $v_2^* = 1$ ,  $v_3^* = 0$  (sect. B.1.4);  $v_1^* = 0$ ,  $0 < v_2^* < 1$ ,  $0 < v_3^* < 1$  (sect. B.1.6);  $v_1^* = 0$ ,  $v_2^* = 1$ ,  $0 < v_3^* < 1$  (sect. B.1.7); and  $v_1^* = 0$ ,  $0 < v_2^* < 1$ ,  $v_3^* = 0$  (sect. B.1.8). In each of the other five cases,

$$v_2^* \leq v_3^*. \quad (\text{B16})$$

As noted in the main body of the article,  $\gamma_c$ , the critical threshold that  $\gamma$  must exceed for  $v^*$  to be an ESS, is a function of  $\lambda$ ,  $\beta$ , and  $r$ . It now follows from equations (B7a) and (B15), section B.1.3, and equations (B10) and (B8a) that this function is defined by

$$\gamma_c(\lambda, \beta, r) = \begin{cases} \frac{1}{2}(1 - \lambda)\lambda^2\{(1 - r)\beta + 2(1 + r)(1 - \lambda)\} & \text{if } (\lambda, \beta) \in \text{A} \\ \frac{1}{2}(1 - \lambda)\{1 + r(1 - \lambda\beta)\}\{\lambda\beta - 2(1 - \lambda)\}/\lambda & \text{if } (\lambda, \beta) \in \text{B} \\ 0 & \text{if } (\lambda, \beta) \in \text{C} \\ \frac{1}{2}(1 + r)\lambda\beta\left\{1 - \frac{\beta}{2(1 - \lambda)}\right\} & \text{if } (\lambda, \beta) \in \text{D} \\ (1 + r)\lambda^2(1 - \lambda)^2 & \text{if } (\lambda, \beta) \in \text{E}, \end{cases} \quad (\text{B17})$$

where the regions A, B, C, D, and E are defined by equations (B7b), (B14), (B9), (B12), and (B8b) and are sketched in figure 3, with the part of region B where  $\lambda\beta < 1$  shaded light in figure 3D, whereas the part where  $\lambda\beta > 1$  is shaded dark. Note that  $\gamma_c$  is at least as large in region A as in region E for any value of  $r$ . For  $v^*$  to be a type IK ESS (for  $K = A, \dots, E$ ), we require both  $(\lambda, \beta)$  to lie in region K and  $\gamma$  to exceed  $\gamma_c(\lambda, \beta, r)$ . Because  $\gamma_c(\lambda, \beta, r)$  varies continuously (and is nonnegative) over the entirety of the semi-infinite rectangle  $0 < \lambda < 1, \beta > 0$  depicted in figure 3, the equation  $\gamma = \gamma_c(\lambda, \beta, r)$  defines a surface, a contour map of which has been superimposed on figure 3 for both  $r = 0$  (noncommunicative assessment) and  $r = 1$  (full communicative assessment). Self-assessment is evolutionarily stable whenever  $\gamma$  exceeds the height of the surface.

## B.2. Mutual Assessment ESS Conditions

For mutual assessment at the ESS, we require  $v_1^* > 0$ . Hence, from equations (A2) and (B1) with  $i = 1$ , we require

$$v_1^* = \frac{1}{(1+r)\alpha} \ln(\zeta), \quad (\text{B18})$$

where  $\zeta$  must not only satisfy the polynomial equation

$$2\gamma\zeta^2 - \lambda(1-\lambda)\{2(1+r)\{\lambda v_2^*(1-v_2^*) + (1-\lambda)v_3^*(1-v_3^*)\} + (1-r)\lambda\beta v_2^*\}\zeta - 2(1+r)(1-\lambda)^2\lambda^2(1-v_2^*-v_3^*)^2 = 0 \quad (\text{B19})$$

but also must exceed 1. For boundary values of  $v_2^*$  and  $v_3^*$  (which are independent of  $\zeta$ ), this is a quadratic equation with precisely one positive root (except in the case where  $v_2^* = 0$  and  $v_3^* = 1$ , which cannot arise; see sect. B.2.3); and  $\zeta > 1$  becomes

$$(1-r)\lambda^2(1-\lambda)\beta v_2^* + 2(1+r)\lambda(1-\lambda)\{v_3^* + \lambda(1-v_2^*-v_3^*)\}\{1-v_3^* - \lambda(1-v_2^*-v_3^*)\} > 2\gamma. \quad (\text{B20})$$

From equation (A2) we obtain

$$\begin{aligned} \frac{2\zeta^2}{\lambda^2 C} \frac{\partial f}{\partial u_2} \Big|_{u=v=v^*} &= (\beta - 2v_2^*)\zeta^2 - (1-\lambda)(\beta - 4v_2^* + 2)\zeta \\ &+ \{(\beta + 2)(1-v_3^*) - 2v_2^*\}(1-\lambda)^2 \end{aligned} \quad (\text{B21})$$

and

$$\begin{aligned} \frac{2\zeta^2}{(1-\lambda)^2 C} \frac{\partial f}{\partial u_3} \Big|_{u=v=v^*} &= (\beta - 2v_3^*)\zeta^2 - 2\lambda(1-2v_3^*)\zeta \\ &+ \{(\beta - 2)v_2^* + 2(1-v_3^*)\}\lambda^2, \end{aligned} \quad (\text{B22})$$

where  $\zeta$  is defined by equation (B19); also

$$\begin{aligned} \frac{2\zeta^2}{C} \frac{\partial^2 f}{\partial u_1^2} \Big|_{u=v=v^*} &= -\alpha\lambda(1-\lambda)\{2(1+r)^2\lambda(1-\lambda)(1-v_2^*-v_3^*)^2 \\ &+ \{(1-r^2)\lambda\beta v_2^* + 2(1+r^2)\{(1-\lambda)(1-v_3^*)v_3^* + \lambda(1-v_2^*)v_2^*\}\}\zeta\}, \end{aligned} \quad (\text{B23})$$

so that  $\partial^2 f / \partial u_1^2|_{u=v=v^*}$  is always negative.

We now consider each of the nine possibilities with  $v_1^* > 0$  in turn.

### B.2.1. Type IIa ESS with $v_1^* > 0, v_2^* = 1, v_3^* = 1$

Here, from equation (B19),  $\zeta$  is the only positive root of the quadratic equation

$$2\gamma\zeta^2 - \lambda^2(1-\lambda)(1-r)\beta\zeta - 2(1+r)(1-\lambda)^2\lambda^2 = 0, \quad (\text{B24})$$

and  $\zeta > 1$  requires

$$\lambda^2(1-\lambda)\{(1-r)\beta + 2(1+r)(1-\lambda)\} > 2\gamma \quad (\text{B25})$$

by equation (B20). From equations (B3) with  $i = 2$  and (B21), we require in addition that

$$(\beta - 2)(\zeta - 1 + \lambda)\zeta \geq 2(1 - \lambda)^2, \quad (\text{B26})$$

where  $\zeta$  is defined by equation (B24), and hence in particular that  $\beta > 2$ , corresponding to the uppermost shaded and unshaded regions of figure B5. Then, from equation (B22), equation (B3) holds with strict inequality for  $i = 3$  because  $\beta > 2$  implies  $(\beta - 2)(\zeta^2 + \lambda^2) + 2\lambda\zeta > 0$ . We will refer to such an ESS as type IIa. If equation (B26) holds with strict inequality, then  $v^*$  is a strong ESS.

*B.2.2. Type IIe ESS with  $v_1^* > 0$ ,  $v_2^* = 0$ ,  $v_3^* = 0$*

Here, from equation (B20), we require

$$(1 + r)\lambda^2(1 - \lambda)^2 > \gamma \quad (\text{B27})$$

or  $\gamma < \gamma_E$ , in which case

$$\zeta = \lambda(1 - \lambda)\sqrt{\frac{(1 + r)}{\gamma}} \quad (\text{B28})$$

by equation (B19). From equations (B3) with  $i = 2, 3$  and (B21) and (B22), we require in addition that

$$\beta\zeta^2 - (1 - \lambda)(\beta + 2)\zeta - (\beta + 2)(1 - \lambda)^2 \leq 0, \quad (\text{B29a})$$

$$(\beta\zeta - 2\lambda)\zeta + 2\lambda^2 \leq 0, \quad (\text{B29b})$$

the second of which implies the first. Note that because  $\zeta > 1$ , equation (B29b) implies  $\beta < 2\lambda$ . Furthermore, equation (B29b) implies  $\beta \leq 2\lambda(\zeta - \lambda)/\zeta^2$ . But  $2\lambda(\zeta - \lambda)$  has an absolute maximum of  $(1/2)\zeta^2$  for  $\lambda = (1/2)\zeta$ , and so it also follows that  $\beta \leq 1/2$ , which restricts this case to the dark shaded region of figure B5. (The left-hand boundary of this region is actually an arc of the curve  $\lambda\beta^3 + 4(\beta - 2\lambda) = 0$ , as opposed to the line  $\beta = 2\lambda$ ; however, the curve and the line are virtually indistinguishable for  $\lambda \leq 1/4$ , with the line lying ever so slightly above the curve.) We will refer to such an ESS as type IIe. If equation (B29) holds with strict inequality, then  $v^*$  is a strong ESS.

*B.2.3. ESS with  $v_1^* > 0$ ,  $v_2^* = 0$ ,  $v_3^* = 1$*

Here equation (B20) reduces to  $0 > 2\gamma$ , which is impossible. Hence this case does not arise.

*B.2.4. ESS with  $v_1^* > 0$ ,  $v_2^* = 1$ ,  $v_3^* = 0$*

Here, from equation (B20), we require

$$\lambda^2(1 - \lambda)(1 - r)\beta > 2\gamma, \quad (\text{B30})$$

in which case  $\zeta$  is the only positive root of (B19) or

$$\zeta = \frac{\lambda^2(1 - \lambda)(1 - r)\beta}{2\gamma}. \quad (\text{B31})$$

From equations (B3) with  $i = 2$ , (B2) with  $i = 3$ , and (B21) and (B22)) we require in addition that

$$(\beta - 2)\zeta^2 - (1 - \lambda)(\beta - 2)\zeta + \beta(1 - \lambda)^2 \geq 0, \quad (\text{B32a})$$

$$\beta\zeta^2 - 2\lambda\zeta + \beta\lambda^2 \leq 0, \quad (\text{B32b})$$

and hence, on multiplying the first of these inequalities by  $-1$  before adding them, that  $\beta\zeta(1 - \lambda) + \beta(2\lambda - 1) + 2\zeta(\zeta - 1) \leq 0$ . However, this inequality can never hold because  $\zeta > 1$ , and so the left-hand side must exceed  $\beta(1 - \lambda) + \beta(2\lambda - 1) + 2\zeta(\zeta - 1) = \beta\lambda + 2\zeta(\zeta - 1)$ , which is positive. So this case does not arise.

*B.2.5. Type IIId ESS with  $v_1^* > 0$ ,  $v_2^* = 0$ ,  $0 < v_3^* < 1$*

Here, with  $v_2^* = 0$ , equation (B1) must hold for  $i = 1, 3$  and equation (B2) must hold for  $i = 2$ . Thus, with  $v_2^* = 0$ , the right-hand side of (B22) must equal 0, yielding

$$v_3^* = \frac{\beta\zeta^2 - 2\lambda\zeta + 2\lambda^2}{2(\zeta - \lambda)^2}, \quad (\text{B33})$$

equation (B19) must hold, which yields the cubic equation

$$(1 + r)\beta\lambda(1 - \lambda)^2\{(\beta - 2)\zeta + 2\lambda\}\zeta + 4\gamma(\zeta - \lambda)^3 = 0 \quad (\text{B34})$$

for  $\zeta$ , on substituting equation (B33) into equation (B19); and from equation (B21),  $\partial f/\partial u_2|_{u=v=v^*} \leq 0$  requires

$$2\beta(\zeta - 1)\{\zeta(\zeta - \lambda) + \lambda(1 - \lambda)\} \leq (1 - \lambda)\{\beta^2(1 - \lambda)\zeta + 4(\zeta - 1)(\zeta - \lambda)\}. \quad (\text{B35})$$

Note that equation (B34) implies  $(\beta - 2)\zeta + 2\lambda < 0$  (because  $\zeta > \lambda$ ), hence  $\beta < 2$  and  $\zeta > 2\lambda/(2 - \beta)$ . Also, the left-hand side of equation (B34) must be less than the right-hand side evaluated for  $\beta = 2$ , or

$$\beta < \frac{2(1 - \lambda)}{\zeta - 1} \left\{ \frac{(\zeta - \lambda)^2 + \lambda(1 - \lambda)}{\zeta(\zeta - \lambda) + \lambda(1 - \lambda)} \right\}. \quad (\text{B36})$$

The term in braces cannot exceed 1. Hence if  $\zeta \geq 2$ , then  $\beta < 2(1 - \lambda)$ ; and if  $1 < \zeta < 2$ , then  $2 > 2\lambda/(2 - \beta)$  implies  $\beta < 2 - \lambda$ . Either way,  $\beta < 2 - \lambda$ . We will refer to such an ESS as type IId. For example, with  $r = 0$ ,  $\lambda = 0.5$ ,  $\beta = 0.7$ , and  $\gamma = 0.04$ , we obtain  $\alpha v_1^* = \ln(1.2613) = 0.232$ ,  $v_2^* = 0$ , and  $v_3^* = 0.304$ .

Even if equation (B35) holds with strict inequality,  $v^*$  is only a weak ESS because there are always alternative best replies. That is, with  $u = (v_1^*, v_2^*, u_3)$  we have  $f(u, v^*) = f(v^*, v^*)$  but  $f(v^*, u) - f(u, u) > 0$  for all  $u_2 \neq v_2^*$ . This condition is hard to establish analytically, but it is easy to verify numerically in any specific case. For example, it is verified by figure B1 for the example given above.

#### B.2.6. Type IIf ESS with $v_1^* > 0$ , $0 < v_2^* < 1$ , $0 < v_3^* < 1$

Here equation (B1) must hold for  $i = 2$  and  $i = 3$  as well as for  $i = 1$ . If  $\beta \geq 2$ , then the expression on the right-hand side of equation (B22) must at least equal  $2(1 - v_3^*)(\zeta - \lambda)^2 + 2\zeta\lambda$  and hence be positive, a contradiction; hence  $\beta < 2$ . With  $v_1^*$  satisfying equation (B18), setting the right-hand sides of equations (B21) and (B22) both equal to zero yields a pair of simultaneous linear equations for  $v_2^*$  and  $v_3^*$ , which are readily solved. The resultant expressions for  $v_2^*$  and  $v_3^*$  are quotients of fourth-order polynomials in  $\zeta$ , which also depend on  $\beta$ ,  $\gamma$ ,  $\lambda$ , and  $r$ ; they are both negative if  $\beta = 0$ , indicating that this case does not arise for small positive values of  $\beta$ , but in practice its minimum feasible value can be found only by numerical means. Substituting the expressions for  $v_2^*$  and  $v_3^*$  into equation (B19) yields a ninth-order polynomial equation for  $\zeta$  that is far too cumbersome to present. Nevertheless, it is readily solved by numerical means for any given values of  $r$ ,  $\lambda$ ,  $\beta$ , and  $\gamma$ . For example, with  $r = 0$ ,  $\lambda = 0.2$ ,  $\beta = 0.7$ , and  $\gamma = 0.01$ , there are two solutions greater than 1, but only the larger one corresponds to a positive value of  $v_2^*$ . We then obtain  $\alpha v_1^* = \ln(3.824) = 1.34$ ,  $v_2^* = 0.171$ , and  $v_3^* = 0.334$ . Note that  $v_3^* > v_2^*$ .

We will refer to such an ESS as type IIf. A type-IIf ESS is always a weak ESS, because there are always alternative best replies. That is, with  $u = (v_1^*, u_2, v_3^*)$  or  $u = (v_1^*, v_2^*, u_3)$ , we have  $f(u, v^*) = f(v^*, v^*)$ , but  $f(v^*, u) - f(u, u) > 0$  for all  $u_2 \neq v_2^*$  or all  $u_3 \neq v_3^*$ , respectively. These conditions are verified by figure B2 for the example given above.

#### B.2.7. ESS with $v_1^* > 0$ , $v_2^* = 1$ , $0 < v_3^* < 1$

Here, with  $v_2^* = 1$ , equation (B1) must hold for  $i = 1, 3$ , and equation (B3) must hold for  $i = 2$ . Thus, with  $v_2^* = 1$ , the right-hand side of (B22) must equal 0, yielding

$$v_3^* = \frac{\beta\zeta^2 - 2\lambda\zeta + \beta\lambda^2}{2(\zeta - \lambda)^2}; \quad (\text{B37})$$

equation (B19) must hold, which yields the quintic equation

$$4\gamma\zeta^2(\zeta - \lambda)^3 + \lambda(1 - \lambda)\left\{(1 + r)(1 - \lambda)\{4\lambda\zeta^3 + \beta^2(\zeta^2 + \lambda^2)^2\} - 2\beta\zeta\left\{(1 + r)(\zeta + \lambda)(\zeta^2 + \lambda^2) - 2\lambda(r\zeta^3 + \lambda^3 + \lambda\zeta\{(2 - r)\zeta + (2r - 1)\lambda\})\right\}\right\} = 0 \quad (\text{B38})$$

for  $\zeta$ , on substituting equation (B37) into equation (B19); and the right-hand side of equation (B21) must not be negative, requiring

$$2(\beta - 2)\zeta(\zeta - 1)\{\zeta(\zeta - \lambda) + \lambda(1 - \lambda)\} \geq \beta^2(1 - \lambda)^2\{\zeta^2 + \lambda^2\} \quad (\text{B39})$$

and hence in particular  $\beta > 2$ . From equation (B37), however,  $v_3^* < 1$  requires  $(\beta - 2)(\zeta^2 + \lambda^2) + 2\lambda\zeta < 0$ , and this contradiction ensures that this case cannot arise.

B.2.8. Type IIg ESS with  $v_1^* > 0$ ,  $0 < v_2^* < 1$ ,  $v_3^* = 0$ 

Here, with  $v_3^* = 0$ , equation (B1) must hold for  $i = 1, 2$  and equation (B2) must hold for  $i = 3$ . Thus, with  $v_3^* = 0$ , the right-hand side of equation (B21) must equal 0, yielding

$$v_2^* = \frac{\beta\{\zeta(\zeta - 1 + \lambda) + (1 - \lambda)^2\} - 2(1 - \lambda)(\zeta - 1 + \lambda)}{2(\zeta - 1 + \lambda)^2}, \quad (\text{B40})$$

equation (B19) must hold, which yields the quintic equation

$$4\gamma\zeta^2(\zeta - 1 + \lambda)^3 + \lambda^2(1 - \lambda)\{\beta^2\{\zeta(\zeta - 1 + \lambda) + (1 - \lambda)^2\}\{2r\zeta(\zeta - 1 + \lambda) + (1 + r)(1 - \lambda)^2\} - 2\beta\zeta(\zeta - 1 + \lambda)\{r\zeta^2 + (1 - \lambda)^2 + (\zeta - 1 + \lambda)^2\}\} = 0 \quad (\text{B41})$$

for  $\zeta$ , on substituting equation (B40) into equation (B19); and the right-hand side of equation (B22) must not be positive, requiring

$$\lambda^2\beta^2\{\zeta(\zeta - 1 + \lambda) + (1 - \lambda)^2\} + 2\beta\zeta^2(\zeta - 1)(\zeta - 1 + 2\lambda) \leq 4\lambda\zeta(\zeta - 1)(\zeta - 1 + \lambda). \quad (\text{B42})$$

Because the second term on the left-hand side must be strictly smaller than the right-hand side when the coefficient 2 in the last factor is replaced by 1, it is readily seen that  $\beta\zeta < 2\lambda$ . Because the first term on the left-hand side must be strictly smaller than the right-hand side, we also have  $\lambda\beta^2 < 4(\zeta - 1)$ . Combining these inequalities, we find that the point  $(\lambda, \beta)$  must lie below the curve  $\lambda\beta^3 + 4(\beta - 2\lambda) = 0$ , as indicated in figure B5. Numerical experiment indicates that the above conditions can be satisfied if both  $\gamma$  is sufficiently small and  $\lambda$  is sufficiently large; for example, with  $r = 1$ ,  $\lambda = 0.8$ ,  $\beta = 0.2$ , and  $\gamma = 0.01$ , we obtain  $\alpha v_1^* = \ln(2.518) = 0.4617$ ,  $v_2^* = 0.02309$ , and  $v_3^* = 0$ .

We will refer to such an ESS as type IIg. A type-IIg ESS is always a weak ESS, because there are always alternative best replies. That is, with  $u = (v_1^*, u_2, v_3^*)$ , we have  $f(u, v^*) = f(v^*, v^*)$ , but  $f(v^*, u) - f(u, u) > 0$  for all  $u_2 \neq v_2^*$ . This condition is verified by figure B3 for the example given above.

This is the only type of ESS for which  $v_2^* > v_3^*$ . Note, however, that this case does not arise when  $\lambda = 1/2$  (app. C).

 B2.9. Type IIb ESS with  $v_1^* > 0$ ,  $0 < v_2^* < 1$ ,  $v_3^* = 1$ 

Here, with  $v_3^* = 1$ , equation (B1) must hold for  $i = 1, 2$  and equation (B3) must hold for  $i = 3$ . Thus, with  $v_3^* = 1$ , the right-hand side of (B21) must equal 0, yielding

$$v_2^* = \frac{\zeta\{\beta(\zeta - 1 + \lambda) - 2(1 - \lambda)\}}{2(\zeta - 1 + \lambda)^2}. \quad (\text{B43})$$

For  $v_2^* < 1$  we require

$$\beta < \frac{2\{\zeta(\zeta - 1 + \lambda) + (1 - \lambda)^2\}}{\zeta(\zeta - 1 + \lambda)}, \quad (\text{B44})$$

and because the right-hand side of equation (B44) decreases with  $\zeta$ , it cannot exceed its value as  $\zeta \rightarrow 1$ . We deduce that

$$\beta < \frac{2(1 - \lambda + \lambda^2)}{\lambda}. \quad (\text{B45})$$

Also, equation (B19) must hold, which yields the cubic equation

$$(1 - \lambda)\lambda^2\{2(1 - \lambda) - \beta(\zeta - 1 + \lambda)\}\{(1 + r)\zeta - r\beta(\zeta - 1 + \lambda)\} + 2\gamma(\zeta - 1 + \lambda)^3 = 0 \quad (\text{B46})$$

for  $\zeta$ , on substituting equation (B43) into equation (B19); and the right-hand side of equation (B22) must not be negative, requiring

$$\beta^2\lambda^2(\zeta - 1 + \lambda) + 2\beta\zeta(\zeta - 1)(\zeta - 1 + 2\lambda) \geq 4(\zeta - 1)\{\zeta(\zeta - 1 + \lambda) + \lambda(1 - \lambda)\} \quad (\text{B47})$$

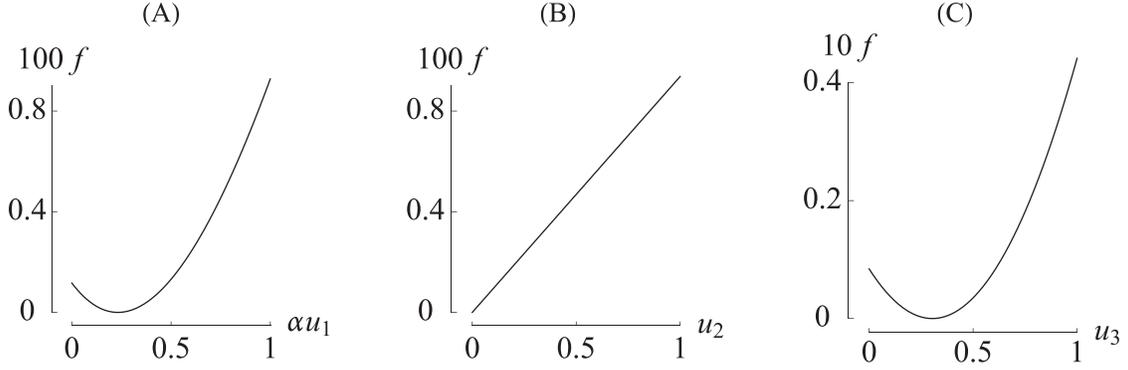
and hence that

$$\beta > \min\left\{1, \frac{2(1 - \lambda)}{1 + \lambda}\right\}. \quad (\text{B48})$$

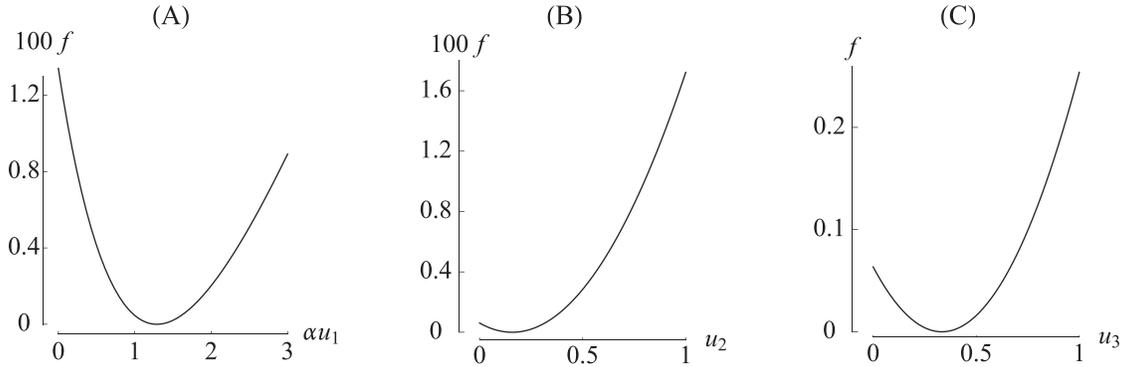
For suppose that  $\beta \leq 1$ . Then the difference between the left- and right-hand sides of equation (B47) cannot exceed  $2\zeta^2(2 - \zeta) - (1 - \lambda)(1 + 5\lambda)\zeta - \zeta + \lambda(1 - \lambda)(4 - \lambda)$ , which is negative if  $\zeta \geq 2$  (because the last term is strictly less than 1); and if  $1 < \zeta < 2$ , then  $v_2^* > 0$  requires  $(1 + \lambda)\beta > 2(1 - \lambda)$ .

We will refer to such an ESS as type IIb. For example, with  $r = 0$ ,  $\lambda = 0.2$ ,  $\beta = 3$ , and  $\gamma = 0.01$ , we obtain  $\alpha v_1^* = \ln(1.3603) = 0.3077$ ,  $v_2^* = 0.1751$ , and  $v_3^* = 1$ . A type-IIb ESS is a weak ESS because there are always alternative best replies. That is, with  $u = (v_1^*, u_2, v_3^*)$  we have  $f(u, v^*) = f(v^*, v^*)$  but  $f(v^*, u) - f(u, u) > 0$  for all  $u_2 \neq v_2^*$ . This condition is verified by figure B4 for the example given above.

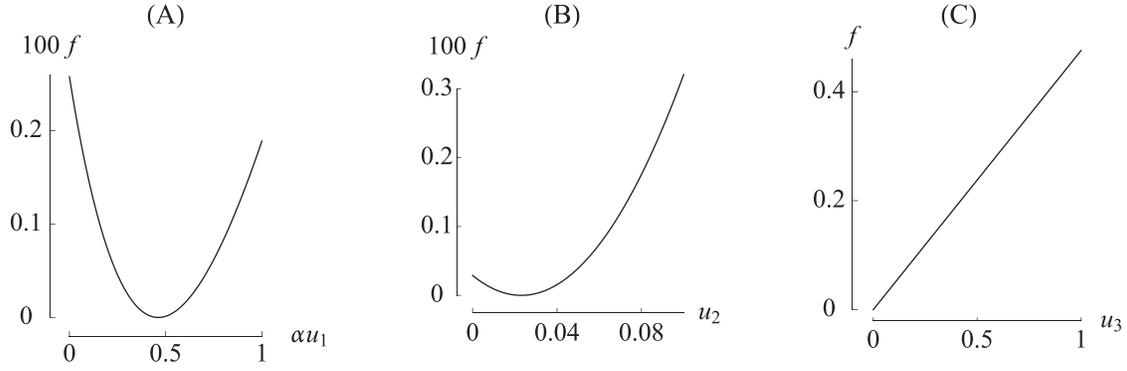
Of the nine cases so far considered in this section, we have now excluded three possibilities, namely,  $v_1^* > 0$ ,  $v_2^* = 0$ , and  $v_3^* = 1$  (sect. B.2.3);  $v_1^* > 0$ ,  $v_2^* = 1$ , and  $v_3^* = 0$  (sect. B.2.4); and  $v_1^* > 0$ ,  $v_2^* = 1$ , and  $0 < v_3^* < 1$  (sect. B.2.7). Constraints on the remaining six possibilities for a type II ESS are summarized by figure B5.



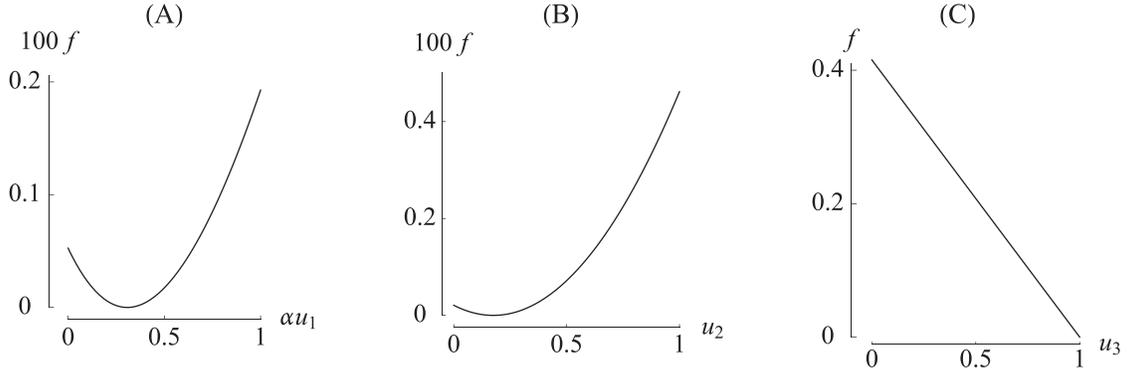
**Figure B1:** Verification of sufficiency conditions for a Type-IIb ESS with  $r = 0$ ,  $\lambda = 0.5$ ,  $\beta = 0.7$ , and  $\gamma = 0.04$ , and hence  $\alpha v_1^* = 0.232$ ,  $v_2^* = 0$ ,  $v_3^* = 0.304$ . A, With  $u = (u_1, v_2^*, v_3^*)$ ,  $f(v^*, v^*) - f(u, v^*)$  is positive for all  $u \neq v^*$ . B, With  $u = (v_1^*, u_2, v_3^*)$ ,  $f(v^*, v^*) - f(u, v^*)$  is positive for all  $u \neq v^*$ . C, With  $u = (v_1^*, v_2^*, u_3)$ ,  $f(v^*, v^*) = f(u, v^*)$  for all  $u$  but  $f(v^*, u) - f(u, u)$  is positive for all  $u \neq v^*$ .



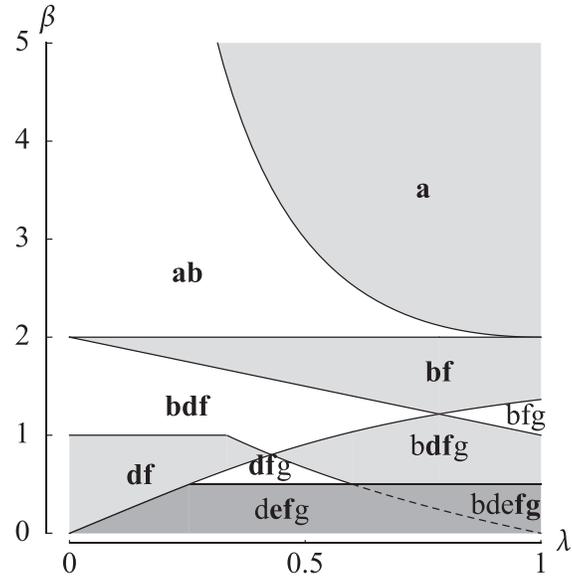
**Figure B2:** Verification of sufficiency conditions for a type-IIc ESS with  $r = 0$ ,  $\lambda = 0.2$ ,  $\beta = 0.7$ , and  $\gamma = 0.01$ , and hence  $\alpha v_1^* = 1.34$ ,  $v_2^* = 0.171$ , and  $v_3^* = 0.334$ . A, With  $u = (u_1, v_2^*, v_3^*)$ ,  $f(v^*, v^*) - f(u, v^*)$  is positive for all  $u \neq v^*$ . B, With  $u = (v_1^*, u_2, v_3^*)$ ,  $f(v^*, v^*) = f(u, v^*)$  for all  $u$  but  $f(v^*, u) - f(u, u)$  is positive for all  $u \neq v^*$ . C, With  $u = (v_1^*, v_2^*, u_3)$ ,  $f(v^*, v^*) = f(u, v^*)$  for all  $u$  but  $f(v^*, u) - f(u, u)$  is positive for all  $u \neq v^*$ .



**Figure B3:** Verification of sufficiency conditions for a type-IIg ESS with  $r = 1$ ,  $\lambda = 0.8$ ,  $\beta = 0.2$ , and  $\gamma = 0.01$ , and hence  $\alpha v_1^* = 0.4617$ ,  $v_2^* = 0.02309$ ,  $v_3^* = 0$ . *A*, With  $u = (u_1, v_2^*, v_3^*)$ ,  $f(v^*, v^*) - f(u, v^*)$  is positive for all  $u \neq v^*$ . *B*, With  $u = (v_1^*, u_2, v_3^*)$ ,  $f(v^*, v^*) = f(u, v^*)$  for all  $u$  but  $f(v^*, u) - f(u, u)$  is positive for all  $u \neq v^*$ . *C*, With  $u = (v_1^*, v_2^*, u_3)$ ,  $f(v^*, v^*) - f(u, v^*)$  is positive for all  $u \neq v^*$ .



**Figure B4:** Verification of sufficiency conditions for a type-IIb ESS with  $r = 0$ ,  $\lambda = 0.2$ ,  $\beta = 3$ , and  $\gamma = 0.01$ , and hence  $\alpha v_1^* = 0.3077$ ,  $v_2^* = 0.1751$ , and  $v_3^* = 1$ . *A*, With  $u = (u_1, v_2^*, v_3^*)$ ,  $f(v^*, v^*) - f(u, v^*)$  is positive for all  $u \neq v^*$ . *B*, With  $u = (v_1^*, u_2, v_3^*)$ ,  $f(v^*, v^*) = f(u, v^*)$  for all  $u$  but  $f(v^*, u) - f(u, u)$  is positive for all  $u \neq v^*$ . *C*, With  $u = (v_1^*, v_2^*, u_3)$ ,  $f(v^*, v^*) - f(u, v^*)$  is positive for all  $u \neq v^*$ .



**Figure B5:** Regions of the  $\lambda - \beta$  plane corresponding to analytical constraints on the different types of mutual-assessment ESS. The diagram indicates regions from which a type IIk ESS—defined for  $k = a, b, d, e, f, g$  in table 2—has not been excluded analytically, with boldface indicating that the ESS type has been found somewhere in that region (even if only at one extreme). Thus, for example, a type IIa ESS may correspond either to the shaded region a or to the unshaded region ab, but only a type IIa ESS can correspond to region a; likewise, a type IIe ESS can be found only in region e, from which a type IIb ESS has been excluded only to the left of the dashed curve.

# Appendix C from M. Mesterton-Gibbons and S. M. Heap, “Variation between Self- and Mutual Assessment in Animal Contests” (Am. Nat., vol. 183, no. 2, p. 199)

## The Case of a Maximum-Variance Population

The results of section B.2 are quite complicated and thus arguably somewhat at odds with our overall goal of using a minimal model to clarify central issues. Accordingly, we specialize to  $\lambda = 1/2$ , which simplifies matters considerably. Now contestants are equally likely to be matched or unmatched, and so variance of opponents is at a maximum.

The critical threshold defined by equation (B17) now depends only on  $\beta$  and  $r$ , and so for any fixed  $r$ , we can clearly identify all five categories of type-I ESS in a  $\beta$ - $\gamma$  plane. The result is shown for  $r = 0$  and  $r = 1$  in figure 4, where there is always a type-IK ESS in shaded region K (for  $K = A, \dots, E$ ). For intermediate values of  $r$ , the corresponding regions are intermediate in the sense that increasing  $r$  from 0 to 1 raises the base of region E and lowers the base of region A while maintaining the contiguity of all regions. In the unshaded regions, only a type II ESS can arise. Above the dashed curves in the shaded regions, only a type I ESS can arise. Below the dashed curves in the shaded regions, however, a type II ESS can arise alongside the type I ESS for that region. Thus, the ESS is not in general unique. For example, at the point  $(\beta, \gamma) = (1.4, 0.01)$  in figure 5A for  $r = 0$ , there is both a type IIb ESS with  $\alpha v_1^* = 0.2257$ ,  $v_2^* = 0.0603$ , and  $v_3^* = 1$  and a type IIc ESS with  $\alpha v_1^* = 2.425$ ,  $v_2^* = 0.685$ , and  $v_3^* = 0.7196$ .

To appreciate these points more fully, we first consider a type IIa ESS. From section B.2.1, a type-IIa ESS arises only if  $\beta > 2$ . From equation (B24), for  $\lambda = 1/2$ , the ESS arises with

$$\zeta = \frac{(1-r)\beta + \sqrt{(1-r)^2\beta^2 + 64(1+r)\gamma}}{32\gamma} \quad (C1)$$

whenever equations (B25) and (B26) are satisfied with strict inequality. From equation (B25), we require the point  $(\beta, \gamma)$  in figure 4 to lie below the line with equation  $\gamma = (1/16)\{1 + r + (1-r)\beta\}$ ; and from equation (B26), we require  $(\beta - 2)\zeta(2\zeta - 1) > 1$ , which must hold for  $\beta > 3$  but for  $2 < \beta < 3$  requires

$$\gamma < \frac{1}{16} \{(\beta - 2)\{1 + (1 + \beta)r\} + |1 - (\beta - 1)r|\sqrt{(\beta - 2)(\beta + 6)}\} \quad (C2)$$

as well as  $\gamma < (1/16)\{1 + r + (1-r)\beta\}$ . In figure 5, equation (C2) demarcates the left-hand boundary of the dark shaded region marked ab. In the slightly darker region to the right, the ESS is uniquely type IIa. Comparing with figure 4, we see that the region in which a type IIa ESS arises does not intersect region A, but it overlaps the lowest part of region B, as indicated by the dashed lines.

Next we consider a type IIe ESS. From section B.2.2, a type IIe ESS arises only if  $\beta \leq 1/2$ . The ESS arises if equation (B28) satisfies equations (B27) and equation (B29b) with strict inequality, that is, if

$$\frac{(1+r)\beta^2}{8(1-\beta) + \sqrt{1-2\beta}} < \gamma < \frac{1+r}{16}. \quad (C3)$$

The corresponding region is shown in figure 5.

For  $\lambda = 1/2$  a type-IIg ESS (sect. B.2.8) does not exist. To see this, we note that the second term on the left-hand side of equation (B42) must be strictly smaller than the right-hand side. For  $\lambda = 1/2$ , the difference reduces to  $\zeta(\zeta - 1)(2\beta\zeta^2 - 2\zeta + 1)$ , which must be negative, and so  $2\beta\zeta^2 - 2\zeta + 1 < 0$ . But then the numerator of equation (B40), which reduces to  $(1/2)(2\beta\zeta^2 - 2\zeta + 1) + (1/4)(1 - 2\zeta)\beta$ , is also negative, contradicting  $v_2^* > 0$ .

It is straightforward to see that, for  $\lambda = 1/2$ , a type-IIb ESS (sect. B.2.9) can exist only if  $1 < \beta < 3$ . We infer  $\beta < 3$  from equation (B45). To show that  $\beta > 1$ , suppose that it is possible to have  $\beta \leq 1$ . Then, because the left-hand side of equation (B47) increases with  $\beta$ , the inequality can hold only if it holds for  $\beta = 1$ , that is, if  $7/8 - (11/4)\zeta + 4\zeta^2 - 2\zeta^3 \geq 0$ , which requires  $\zeta \leq 1.12208$ . From equation (B43),  $v_2^* > 0$  requires  $\zeta > 1/2 + 1/\beta \geq 1.5$ , and from this contradiction we infer that  $\beta \leq 1$  is false. But 1 is not the greatest lower bound for  $\beta$ . A more refined argument, based on the discussion that follows, shows that a type-IIb ESS exists only if  $(5)^{1/2} - 1 < \beta < 3$  (regardless of the value of  $r$ ).

A type-IIb ESS will actually exist for  $(5)^{1/2} - 1 < \beta < 3$ , however, only if  $\gamma$  lies in a suitable range; outside this range, the ESS is of a different type. To obtain upper and lower bounds on  $\gamma$ , we first rewrite (B46) for  $\lambda = 1/2$  as

$$\gamma = \frac{\{\beta(2\zeta - 1) - 2\}\{2(1 + r)\zeta - r\beta(2\zeta - 1)\}}{8(2\zeta - 1)^3}. \quad (\text{C4})$$

Note that the right-hand side of this equation is always positive when  $\zeta > 1/2 + 1/\beta$  (required by  $v_2^* > 0$ ) if  $\beta \leq 1 + 1/r$ , but is negative when  $\zeta > r\beta/\{2(r\beta - r - 1)\} > 1/2 + 1/\beta$  if  $\beta > 1 + 1/r$ , which can happen only if  $r > 1/2$  (because  $\beta < 3$ ). A type-IIb ESS cannot exist for any value of  $\gamma$  that either exceeds the maximum of the right-hand side of equation (C4) or fails to exceed its minimum, which can be shown to require  $\gamma > \max\{0, (1/2)[(1/2)\beta - 1]\}$  if  $r = 0$  but to require  $\gamma > \max\{0, (1/16)(\beta - 2)(\beta + 2 - [(\beta - 2)(\beta + 6)]^{1/2})\}$  if  $r = 1$ ; however, the corresponding expression for intermediate values of  $r$  is too unwieldy to be useful, even though it is straightforward to find analytically. Moreover, any attempt to locate the upper bound analytically likewise leads to excessively unwieldy algebra that is scarcely conducive to clarity. Instead, therefore, we have found the upper limit of  $\gamma$ 's range by numerical optimization and presented the results in figure 5. Note that this optimization is constrained (possible values of  $\zeta$  must satisfy  $\zeta > 1$ ,  $0 < v_2^* < 1$ , and  $\partial f/\partial u_3|_{u=v=v^*} > 0$ ).

A similar analysis applies to a type-IId ESS, which (for  $\lambda = 1/2$ ) can exist only if  $\beta < 3/2$  (sect. B.2.5), and to a type-IIf ESS, which can exist only if  $\beta < 2$  (sect. B.2.6). To obtain upper and lower bounds for a type-IId ESS, we first rewrite equation (B34) as

$$\gamma = \frac{(1 + r)\beta\zeta\{(2 - \beta)\zeta - 1\}}{4(2\zeta - 1)^3} \quad (\text{C5})$$

and then optimize it numerically, as described above, to yield the results for  $r = 0$  and  $r = 1$  that are plotted in figure 5. Likewise for type IIf, although we refrain from presenting the corresponding analogue of equation (C5); as noted in section B.2.5, it is excessively unwieldy.

**Appendix D from M. Mesterton-Gibbons and S. M. Heap, “Variation between Self- and Mutual Assessment in Animal Contests” (Am. Nat., vol. 183, no. 2, p. 199)**

**Variation in Withdrawal Strategy and Alternative ESS Conditions**

Analysis of the model also indicates that variation in withdrawal strategy is associated with variation in assessment strategy (table 2). First, some withdrawal strategies are dependent on a particular mode of assessment. In particular, unconditional withdrawal or persistence based on perceiving one’s opponent to have the same RHP can be evolutionarily stable only under self-assessment (table 2c). Similarly, only under mutual assessment can the probability of persistence in contests perceived as matched be greater for weak contestants (table 2g) or mixed with withdrawal for both RHP categories (table 2f). Second, the mode of assessment determines limits (in terms of  $\beta$ ,  $\lambda$ , and  $\gamma$ ) for a given withdrawal strategy to be part of the ESS. For example, when  $\lambda = 1/2$  (maximum variance), unconditional hawk (table 2a) requires relative resource value ( $\beta$ ) to exceed 3 with self-assessment (fig. 4) but is possible as long as  $\beta > 2$  with mutual assessment (fig. 5). There are also regions in which either self- or mutual assessment can be evolutionarily stable for the same values for  $\beta$ ,  $\lambda$ , and  $\gamma$ , which is due to the adoption of alternative withdrawal strategies (fig. 5). For instance, even when  $\gamma_c = 0$ , mutual assessment can be evolutionarily stable if either weak individuals are more aggressive or strong individuals are less aggressive than under self-assessment (table 2b, 2d, 2f cf. 2c).