APPENDIX A: Finding the ESS

(a) *Characterising the population of opponents*

The chance, s, that a randomly encountered opponent in the population who plays Hawk is strong is (by Bayes' theorem)

$$s = \frac{p_1 h_1}{p_0 h_0 + p_1 h_1}, \qquad \text{(A1)}$$

where p_1 is the proportion of individuals that are strong, $p_0 = 1 - p_1$ is the proportion of individuals that are weak, h_0 is the chance that a randomly encountered weak opponent plays Hawk and h_1 is the chance that a randomly encountered strong opponent plays Hawk. From the contest structure outlined in the main text, we can thus calculate the chance, x_0, that a weak individual wins a fight against a randomly encountered opponent (who plays Hawk) as

$$x_0 = \frac{1}{2}(1-s) + \gamma s \qquad \text{(A2a)}$$

and the chance, x_1, that a strong individual does so as

$$x_1 = (1-\gamma)(1-s) + \frac{1}{2}s, \qquad \text{(A2b)}$$

where γ is the chance that a weak individual defeats a strong opponent. We can use x_0 and x_1 to calculate the state-dependent probability $Q(f,n)$ that an individual is strong, given that it has won n of its f previous fights:

$$Q(f,n) = \frac{p_1 x_1^n (1-x_1)^{f-n}}{p_1 x_1^n (1-x_1)^{f-n} + p_0 x_0^n (1-x_0)^{f-n}}. \qquad \text{(A3)}$$

Then the chance, $B(f,n)$, that an individual with n victories in f previous fights will defeat a randomly encountered opponent (who plays Hawk) is

$$B(f,n) = Q(f,n)\cdot x_1 + (1 - Q(f,n))\cdot x_0\,. \tag{A4}$$

This expression can be used to derive an individual's best-response strategy in the population in question, as explained below.

(b) *Identifying the error-prone best-response strategy*

We first determined the best response when the mutant individual's memory is 'full', in that it has previously experienced at least F fights and therefore will continue to play Hawk with the same probability $P_H(F,n)$. For this individual, the expected future fitness if it plays Hawk in the current encounter, W_H, is given by

$$\begin{aligned} W_H(F,n) = {} & [p_0(1-h_0) + p_1(1-h_1)]\cdot v \\ & + (p_0 h_0 + p_1 h_1)\cdot[B(F,n)\cdot v - (1 - B(F,n))\cdot c] + (1-d)R \end{aligned} \tag{A5a}$$

where R represents the fitness gains from future rounds (see below). The first term gives the pay-off when its opponent plays Dove, while the second gives the pay-off when the opponent also plays Hawk (and therefore the contest escalates into a physical fight). If, instead, the mutant individual plays Dove, its expected future fitness W_D is

$$W_D(F,n) = [p_0(1-h_0) + p_1(1-h_1)]\frac{v}{2} + (1-d)R \tag{A5b}$$

since it receives nothing if its opponent plays Hawk.

We seek the best-response strategy for this mutant in the current population. In the absence of error, the best response is to choose whichever of the two options, Hawk or

Dove, gives the highest pay-off. However, we assume that behavioural decisions are error-prone, such that (for any values of f and n) the chance of playing Hawk is computed as

$$P_H(f,n)=\frac{1}{1+\exp\left[\frac{1}{\varepsilon}\left(W_D(f,n)-W_H(f,n)\right)\right]} \tag{A6}$$

where ε is a small positive constant setting the frequency of mistakes (for the results shown, we used ε = 0.005). Note that equation (A6) yields a value for $P_H(F,n)$ that is independent of R in equations (A5a) and (A5b). The form of equation (A6) implies that costly mistakes are rare (McNamara *et al.* 1997). Taking this into account, the expected future fitness of a best-response mutant with n wins out of f fights is

$$W(f,n)=P_H(f,n)\cdot W_H(f,n)+\left(1-P_H(f,n)\right)\cdot W_D(f,n). \tag{A7}$$

For $f = F$, the mutant's memory is already full, and the decision to play Hawk or Dove does not affect its fitness in future rounds. Consequently, $R=W(F,n)$, and combining equations (A5)–(A7) we have

$$\begin{aligned} W(F,n) &= P_H(F,n)\cdot W_H(F,n)+\left(1-P_H(F,n)\right)\cdot W_D(F,n) \\ &= P_H(F,n)\begin{pmatrix} \left[p_0(1-h_0)+p_1(1-h_1)\right]\cdot v \\ +\left(p_0h_0+p_1h_1\right)\cdot\left[B(F,n)\cdot v-\left(1-B(F,n)\right)\cdot c\right] \end{pmatrix} \\ &\quad +\left(1-P_H(F,n)\right)\left[p_0(1-h_0)+p_1(1-h_1)\right]\frac{v}{2} \\ &\quad +(1-d)\cdot W(F,n) \end{aligned}$$

$$\Rightarrow W(F,n)=\frac{1}{d}\begin{pmatrix} P_H(F,n)\begin{pmatrix} \left[p_0(1-h_0)+p_1(1-h_1)\right]\cdot v \\ +\left(p_0h_0+p_1h_1\right)\cdot\left[B(F,n)\cdot v-\left(1-B(F,n)\right)\cdot c\right] \end{pmatrix} \\ +\left(1-P_H(F,n)\right)\left[p_0(1-h_0)+p_1(1-h_1)\right]\frac{v}{2} \end{pmatrix} \tag{A8}$$

Having determined the best response for the mutant individual when it has a full memory, we worked backwards from this point to consider the same individual after $F-1$ fights, then $F-2$ fights and so on, ending up with naïve individual for which $f = 0$. The calculations are different because with $f < F$, the decision to play Hawk or Dove can affect the individual's state (i.e. its information variables f and n). The expected future fitness pay-offs are computed as

$$W_H(f,n) = [p_0(1-h_0) + p_1(1-h_1)](v + (1-d)\cdot W(f,n)) + (p_0h_0 + p_1h_1)\begin{bmatrix} B(F,n)(v + (1-d)\cdot W(f+1,n+1)) \\ + (1-B(F,n))(-c + (1-d)\cdot W(f+1,n)) \end{bmatrix} \quad \text{(A9a)}$$

$$W_D(f,n) = [p_0(1-h_0) + p_1(1-h_1)]\frac{v}{2} + (1-d)\cdot W(f,n) \quad \text{(A9b)}$$

where $W(f,n)$ is given by equation (A7). Note that when both the mutant individual and its opponent play Hawk, f is incremented by 1 unit; and if the mutant wins, n is also incremented by 1 unit. For each combination of f and n, we used the function FindRoot in *Mathematica* (Wolfram Research, Inc. 2007) to find numerical solutions to equations (A6), (A7) and (A9) in terms of W_H, W_D, $P_H(f,n)$ and W.

(c) *Calculating the resulting levels of aggression*

If the best-response strategy were adopted by all individuals in the population, this would give rise to new values of h_0 and h_1, which we call h_{0b} and h_{1b}. Thus h_{0b} is the chance that a randomly encountered weak individual in the best-response population plays Hawk, while h_{1b} is the chance that a randomly encountered strong individual in the best-response population plays Hawk. We calculated these as the average probability of playing Hawk for all individuals of that fighting ability (weak or strong), weighted by the frequency of individuals in each state:

$$h_{0b} = \frac{\sum_{f=0}^{F}\sum_{n=0}^{f}\delta(f,n)\cdot(1-Q(f,n))\cdot P_H(f,n)}{p_0} \qquad \text{(A10a)}$$

$$h_{1b} = \frac{\sum_{f=0}^{F}\sum_{n=0}^{f}\delta(f,n)\cdot Q(f,n)\cdot P_H(f,n)}{p_1}, \qquad \text{(A10b)}$$

where $\delta(f,n)$ is the frequency of individuals that have won n of their f previous fights ($0 \leq \delta \leq 1$) and is given by

$$\delta(f,n) = \begin{cases} \delta(f,n)\cdot[1-P_H(f,n)\cdot(p_0h_0+p_1h_1)]+d & \text{for } f=n=0 \quad \text{(A11a)} \\ (1-d)\cdot\begin{bmatrix} \delta(f,n)\cdot[1-P_H(f,n)\cdot(p_0h_0+p_1h_1)] \\ +\delta(f-1,n)\cdot P_H(f-1,n)\cdot(p_0h_0+p_1h_1)\cdot[1-B(f-1,n)] \\ +\delta(f-1,n-1)\cdot P_H(f-1,n-1)\cdot(p_0h_0+p_1h_1)\cdot B(f-1,n-1) \end{bmatrix} & \text{otherwise.} \quad \text{(A11b)} \end{cases}$$

Individuals in state $f = n = 0$ (equation (A11a)) were either already in that state in the preceding round (frequency $\delta(0,0)$) and then did not get into a fight (probability $1-P_H(0,0)\cdot(p_0h_0+p_1h_1)$), or were newly born at the end of that round (frequency d). For individuals with $f > 0$, the calculation is more complicated because these individuals could have had one of three different experiences in the preceding round, reflected by the three terms in equation (A11b). The first term represents individuals that were already in state (f,n) and did not get into a fight in the preceding round; similarly to before, this happens with probability $1-P_H(f,n)\cdot(p_0h_0+p_1h_1)$. The second term represents individuals that were previously in state $(f-1,n)$ and then got into a fight (probability

$P_H(f-1,n)\cdot(p_0h_0+p_1h_1)$) which they lost (probability $1-B(f-1,n)$), so their number of fights was updated by 1 unit. The third term represents individuals that were in state (f − 1,n − 1) and then got into a fight (probability $P_H(f-1,n)\cdot(p_0h_0+p_1h_1)$) which they won (probability $B(f-1,n-1)$), so their number of fights and number of victories were both updated by 1 unit. In all three cases, the chance that these individuals survived to the current round is 1 − d. Starting from a population composed entirely of naïve individuals ($\delta(f,n)=1$ for f = n = 0 and $\delta(f,n)=0$ otherwise), equation (A11) was iterated 100 times for all values of f and n to generate a stable frequency distribution. These stable values of $\delta(f,n)$ were then entered into equation (A10) to obtain h_{0b} and h_{1b}, the levels of aggression in a population playing the best-response strategy.

(d) *Updating the values of* h$_0$ *and* h$_1$

We adjusted the population levels of aggression h_0 and h_1 in the direction of the calculated best-response values h_{0b} and h_{1b} to yield updated values h_0' and h_1', according to the following equation:

$$h_0' = (1-\lambda)h_0 + \lambda h_{0b} \quad \text{(A12a)}$$

$$h_1' = (1-\lambda)h_1 + \lambda h_{1b}, \quad \text{(A12b)}$$

where λ is a constant between 0 and 1 controlling the degree of updating.

(e) *Iterating until convergence*

We repeated steps (b)–(d) until the process converged on a stable solution. The process was halted when a best-response strategy was found for which h_{0b} and h_{1b} differed from h_0 and h_1 by less than 0.000001. This strategy was taken to be the ESS.

References

McNamara, J. M., Webb, J. N., Collins, E. J., Székely, T. & Houston, A. I. 1997 A general technique for computing evolutionarily stable strategies based on errors in decision-making. *J. Theor. Biol.* **189**, 211–225. (doi: 10.1006/jtbi.1997.0511)

Wolfram Research, Inc. 2007 *Mathematica*, Version 6.0. Champaign, Illinois: Wolfram Research, Inc.