



## Research

**Cite this article:** Veller C, Haig D, Nowak MA. 2016 The Trivers–Willard hypothesis: sex ratio or investment? *Proc. R. Soc. B* **283**: 20160126. <http://dx.doi.org/10.1098/rspb.2016.0126>

Received: 19 January 2016

Accepted: 18 April 2016

**Subject Areas:**

theoretical biology, evolution, behaviour

**Keywords:**

parental condition, individual sex ratio, parental care

**Authors for correspondence:**

Carl Veller

e-mail: [carlveller@fas.harvard.edu](mailto:carlveller@fas.harvard.edu)

Martin A. Nowak

e-mail: [martin\\_nowak@harvard.edu](mailto:martin_nowak@harvard.edu)

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2016.0126> or via <http://rspb.royalsocietypublishing.org>.

# The Trivers–Willard hypothesis: sex ratio or investment?

Carl Veller<sup>1,2</sup>, David Haig<sup>1</sup> and Martin A. Nowak<sup>1,2,3</sup>

<sup>1</sup>Department of Organismic and Evolutionary Biology, <sup>2</sup>Program for Evolutionary Dynamics, and <sup>3</sup>Department of Mathematics, Harvard University, MA, USA

The Trivers–Willard hypothesis has commonly been considered to predict two things. First, that a mother in good condition should bias the sex ratio of her offspring towards males (if males exhibit greater variation in reproductive value). Second, that a mother in good condition should invest more per son than per daughter. These two predictions differ empirically, mechanistically and, as we demonstrate here, theoretically too. We construct a simple model of sex allocation that allows simultaneous analysis of both versions of the Trivers–Willard hypothesis. We show that the sex ratio version holds under very general conditions, being valid for a large class of male and female fitness functions. The investment version, on the other hand, is shown to hold only for a small subset of male and female fitness functions. Our results help to make sense of the observation that the sex ratio version is empirically more successful than the investment version.

## 1. Introduction

The Trivers–Willard hypothesis (TWH) [1] predicts that, when one sex exhibits more variation in reproductive value, then mothers in good condition should ‘prefer’ offspring of that sex, while mothers in poor condition should ‘prefer’ offspring of the other sex. (In polygynous species, the sex with more variable reproductive value is usually males; for convenience we shall adopt this convention throughout.) What should be meant by ‘prefer’, however, is not entirely clear. Two definitions can be found in the empirical literature.

One definition is that mothers in good condition should bias their progeny *sex ratios* towards sons, and mothers in poor condition towards daughters. This interpretation is suggested by the title of Trivers and Willard’s original paper, ‘Natural selection of parental ability to vary the sex ratio of offspring’. This ‘sex ratio version’ of the TWH has been the subject of a large empirical literature—well-known examples exist for red deer [2,3], feral horses [4], parasitoid wasps [5] and humans [6,7].

The second definition is that mothers in good condition should bias their *parental care* towards sons, and mothers in poor condition towards daughters. This definition also seems to be implied by Trivers & Willard [1, p. 91]: ‘In species with a long period of [parental investment] after birth of young, one might expect biases in parental behaviour toward offspring of different sex, according to parental condition; parents in better condition would be expected to show a bias toward male offspring.’ Many empirical tests of this ‘investment version’ of the TWH have also been conducted. For example, Fujita *et al.* [8] conclude that the TWH is supported by their finding that mothers of low socio-economic status in agropastoral villages in northern Kenya produce more nutritious milk when breastfeeding daughters than sons, and vice versa for mothers of high socio-economic status. The assumption that biased parental investment is part of the TWH extends far into the empirical literature (see, e.g. [9–18], and the references in these papers) and is explicitly defended in [19].

The two versions of the TWH clearly make very different empirical predictions. A broad observation is that the sex ratio version is empirically more successful than the investment version [20]. On the sex ratio version, for example, of the 37 studies of ungulate mammals surveyed by Sheldon & West [21], 29 found a sex ratio effect in the direction predicted by the TWH. Fifteen of the

37 studies returned significant results ( $p < 0.05$ ); of these 15, 13 were in the direction predicted by the TWH.<sup>1</sup> In a statistical meta-analysis of these studies, Sheldon & West [21] found a significant positive correlation between maternal condition and the proportion of male offspring. In humans, adequately powered empirical tests tend to find a small sex ratio bias in the direction predicted by the TWH [24,25].

Empirical tests of the investment version, on the other hand, tend to show more mixed results. Only four of the eight studies of non-human mammals surveyed by Keller *et al.* [20] found an investment effect in the direction predicted by the TWH, and studies in humans have proved similarly inconclusive [20]. There could be methodological reasons for this. For example, while it is obvious what outcome to measure when testing for biased sex ratios, it is not as obvious when testing for biases in parental care, as parental care can take many forms [20]. There also does not yet exist an extensive review of empirical tests of the investment version, and so any statements about its empirical success must be preliminary. With these caveats in mind, it is still safe to conclude that the sex ratio version has thus far been empirically more successful.

The two versions of the TWH also differ mechanistically, an observation which has significant implications for research into the physiological and behavioural bases of the parental biases predicted by the hypothesis. The aim of the present paper is to characterize how the two versions differ in theory. This issue has been raised by Carranza [26], who gave a simple graphical example for which the sex ratio version holds while the investment version does not—though see [19] for a critical review of this example.

Although much theoretical work has been carried out on the TWH and sex allocation, in general, it has typically not focused on the distinction between the two versions of the hypothesis that are tested in the empirical literature. Most models of sex allocation (e.g. [27, §§1–3]) combine sex ratio and investment costs into a single variable for each sex, ‘parental expenditure’ (or ‘allocation’). This variable is useful for a number of general arguments, most notably Fisher’s principle [28,29] that, under certain assumptions, total population parental expenditure on offspring of the two sexes should be equal. Other models of sex allocation have held brood composition fixed, and derived conditions under which parental investment should be biased towards male or female offspring (e.g. [27,§5], [30,31]). Others have focused simultaneously on sex ratio and clutch size in cases where sibling interaction is relevant, but have not studied the level of parental investment/care in individual offspring [32–35].

A unified model allowing for simultaneous analysis of the sex ratio and investment versions of the TWH is therefore lacking. Here, we construct a simple model to expose the logic underpinning both the sex ratio and investment versions of the TWH. We show that, while the sex ratio version holds under very general specifications of how adult condition translates to fitness for males and females, the investment version is very sensitive to the nature of these condition-to-fitness functions. This theoretical result helps to make sense of the disparity in empirical success between the two versions of the hypothesis.

## 2. The general model

The TWH is based on three assumptions. First, that parental condition (usually taken to be mother’s) correlates with

offspring condition. Second, that offspring condition persists into adulthood, so that it is relevant for reproductive value. Third, that ‘Adult males will be differentially helped in [reproductive success] (compared with adult females) by slight advantages in condition’ [1, p. 91].

‘Condition’ is not explicitly defined in the original paper of Trivers & Willard [1]. Its definition has remained vague in much of the biological literature [36], but in the context of the TWH, can involve many physical factors, including the ability of a parent to provision its offspring, as well as heritable factors such as traits under direct genetic control and social status. It is important to note that condition cannot be interpreted simply as reproductive value—either of parent or offspring—because the TWH requires that condition translate to reproductive value differently for males and females.

We capture the three assumptions of the TWH in the following simple model.

Each mother is of a certain physical condition,  $c_m$  (continuously distributed), and has a single brood of precisely two offspring. A mother’s total investment capability, which she apportions fully between her two offspring, is limited by her physical condition to  $I(c_m)$ , assumed to be non-decreasing in  $c_m$  (mothers in better condition are able to invest more) and smooth. The post-investment condition of an offspring whose mother was of condition  $c_m$  and who received investment  $i$  is, regardless of its sex,  $c(c_m, i)$ , increasing in both arguments ( $(\partial/\partial c_m)c(c_m, i) > 0$  and  $(\partial/\partial i)c(c_m, i) > 0$ ) and concave in investment ( $(\partial^2/\partial i^2)c(c_m, i) < 0$ ). This post-investment condition is assumed to persist into adulthood, and to be the quantity that is relevant for offspring fitness.

That  $c$  should depend on  $i$  is clear. In our model, ‘investment’ is taken to mean the level of direct parental care for a single offspring. It is plain that, by this definition, offspring condition should increase with investment.

Why does  $c$  depend directly on  $c_m$  too? Parental condition can affect offspring’s condition directly (as opposed to indirectly, through greater investment capacity) in several ways, examples of which are: (i) if condition depends on social rank, and rank is hereditary—as, for example, in baboons [37], spotted hyenas [38] and red grouse [39]; and (ii) if condition is in part determined genetically (this could apply either to maternal or paternal condition: in our model,  $c_m$  could equally well refer to either). This genetic mechanism of condition transmission is especially clear for sexually selected traits [40–43].<sup>2</sup>

The model thus far captures the first two assumptions of Trivers & Willard [1]. Their third assumption, in the context of our model, means at the very least that adult condition  $c(c_m, i)$  translates differently into reproductive success for males than it does for females. (In fact, the focal quantity should be reproductive value, not reproductive success [22]. We abbreviate this to ‘fitness’.) A son in condition  $c$  will be assumed to have fitness  $f_{\sigma}(c)$ , while a daughter in the same condition will have fitness  $f_{\varphi}(c)$ .

The optimization problem for a mother in condition  $c_m$  is one of backwards induction: first, to choose the optimal investment decision for each possible brood (male–male, male–female and female–female), and then, using the maximum fitness sums that arise from these investment decisions, to choose the sex ratio that maximizes the expected sum of her offspring’s fitnesses. Denoting by  $p$  the probability of having a son (the same probability is assumed to apply independently for

each of the two offspring), the optimization problem can be formally written:

$$\begin{aligned} \max_{p \in [0,1]} \quad & \max_{i_1, i_2, i_3 \in [0, I(c_m)]} p^2 [f_{\sigma}(i_1) + f_{\sigma}(I(c_m) - i_1)] \\ & + 2p(1-p) [f_{\sigma}(i_2) + f_{\varphi}(I(c_m) - i_2)] \\ & + (1-p)^2 [f_{\varphi}(i_3) + f_{\varphi}(I(c_m) - i_3)], \end{aligned}$$

where  $f_{\sigma}(i_1)$ , for example, is shorthand for  $f_{\sigma}(c(c_{mv}, i_1))$ .

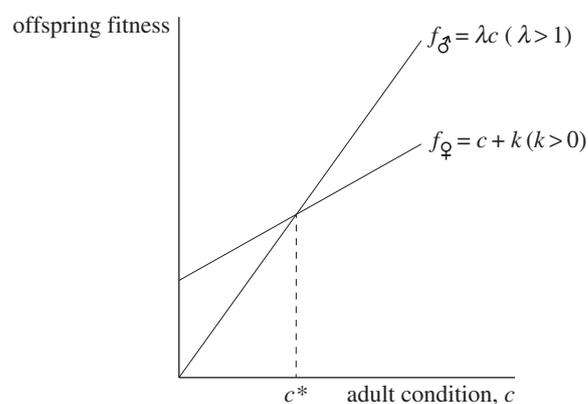
In the context of our model, the sex ratio version of the TWH may be stated as follows: there exists a threshold maternal condition such that a mother in better condition has an optimal progeny sex ratio that is male-biased, while a mother in worse condition has an optimal sex ratio that is female-biased. The question of whether a mother should have a male- or female-biased sex ratio is shown in the electronic supplementary material, appendix S1 to be equivalent to the simpler question of whether she would prefer to have two male or two female offspring (which is not to say that her optimal sex ratio is then necessarily one or zero—this is just a ‘litmus test’ for the optimal sex ratio being biased by any amount).

The investment version of the TWH, on the other hand, is this: there exists a threshold maternal condition such that a mother in better condition, upon having a mixed brood (one male and one female offspring), will invest more in her son than in her daughter, while a mother in worse condition will invest more in her daughter than in her son.

It should be noted, in the light of the facts that checking the sex ratio version amounts to comparing the fitness sum of an all-male brood with that of an all-female brood, and checking the investment version amounts to comparing investment patterns in a mixed brood, that our results are based on simultaneous optimization, by backwards induction, of the sex ratio and investment decisions. In particular, the optimal sex ratio will itself depend on the optimal investment patterns for the various possible brood compositions, as is made clear in electronic supplementary material, appendix S1.

The model presented above is not meant to represent reality in full detail, but rather to make accessible the logic underpinning the two versions of the TWH. More complex models that extend our framework to more realistic settings would also be desirable. For example, we have ignored possible demographic differences between the sexes, though these have recently been shown to influence predictions of the TWH [23]. Our assumption that each mother has exactly two offspring is not realistic but simplifies the analysis considerably, and is the smallest brood number that still allows for simultaneous analysis of sex ratio and investment bias. It is argued in the Discussion section and electronic supplementary material, appendix S3 that the case of a monogamous species, where single offspring are raised in sequence, will not differ much from our simple single-period, two-offspring model.

It will also be noted that our model applies best to contexts where: (i) the sex ratio and parental care decisions are separable (e.g. in time), and (ii) parental care is of a continuous, quantitative form (e.g. the length of the period of parental care—duration of gestation, or time until weaning, say—or the fat content of breast milk). In such cases, the conceptual difference between the sex ratio and investment versions of



**Figure 1.** The ‘classic’ Trivers–Willard fitness functions, for which the sex ratio version of the Trivers–Willard hypothesis holds, while the investment version does not. Because male fitness always increases faster with condition than does female fitness, fitness returns to investment in a male offspring are always greater than in a female offspring of the same initial condition, and so a mother with a mixed brood will always invest more in the male of the brood, no matter her own condition.

the TWH is clear: (i) excludes cases where the sex ratio and investment decisions coincide, such as in parasitoid wasps. It applies, on the other hand, to vertebrate species where parental care is given post-conception and post-birth; and (ii) excludes those discrete parental care decisions that are harder, conceptually, to distinguish from sex ratio decisions, such as sex-biased brood reduction [30,44].

### 3. An example where the investment version fails

A simple case that satisfies Trivers and Willard’s assumption 3, and which results in a mother always investing more in the male of a mixed brood, no matter her condition (and thus contra the investment version of the TWH), is where the fitness-condition profiles are linear for both males and females:  $f_{\sigma}(c) = \lambda c$  and  $f_{\varphi}(c) = c + k$ , with  $\lambda > 1$  so that the fitness profile is steeper for males than for females (clearly satisfying Trivers and Willard’s assumption 3; figure 1).  $k > 0$  can be adjusted arbitrarily; for example, it can be adjusted so that the average reproductive value of sons and daughters coincides. Alternatively, if offspring of one sex cost more to raise than offspring of the other sex, this could be represented by the fitness function of the more expensive sex shifting downwards for each condition; such a shift would lead to a population sex ratio biased against that sex. These facts concerning the adjustability of the fitness functions—i.e. the average fitness effects of having male and female offspring and, therefore, also the population sex ratio—are also true of all later specifications.

The investment decision of a female in condition  $c_m$  who has a mixed brood, denoting by  $i_s$  how much she invests in her son, amounts to the following maximization problem:

$$\begin{aligned} i_s^* &:= \arg \max_{i_s} f_{\sigma}(c(c_m, i_s)) + f_{\varphi}(c(c_m, I(c_m) - i_s)) \\ &= \arg \max_{i_s} [\lambda c(c_m, i_s)] + [c(c_m, I(c_m) - i_s) + k]. \end{aligned}$$

Her investment will be male-biased if  $i_s^* > I(c_m)/2$ , i.e. if  $i_s^* > I(c_m) - i_s^*$ . Assuming an interior solution  $i_s^* \in (0, I(c_m))$ ,

the first-order condition for  $i_s^*$  amounts to equating the male and female offspring's marginal fitness returns to investment:

$$\begin{aligned} \frac{\partial}{\partial i_s} ([\lambda c(c_m, i_s)] + [c(c_m, I(c_m) - i_s) + k]) \Big|_{i_s=i_s^*} &= 0 \\ \Rightarrow \lambda \partial_2 c(c_m, i_s^*) - \partial_2 c(c_m, I(c_m) - i_s^*) &= 0 \\ \Rightarrow \partial_2 c(c_m, i_s^*) < \partial_2 c(c_m, I(c_m) - i_s^*) & \\ \Rightarrow i_s^* > I(c_m) - i_s^* & \end{aligned}$$

where  $\partial_2 c(\cdot, \cdot)$  refers to the partial derivative of  $c$  with respect to its second argument, and where the second-to-last step follows from  $\lambda > 1$  and the last step from the concavity assumption  $\partial_{22} c < 0$  ( $\partial_{22} c(\cdot, \cdot)$  is the second partial derivative). Thus, no matter a mother's condition, she always biases investment in favour of the male offspring of a mixed brood. The reason for this is that the marginal returns to investment in a male offspring,  $\lambda \partial_2 c(c_m, i)$ , are always greater than the returns to investment in a female of the same condition,  $\partial_2 c(c_m, i)$ .<sup>3</sup>

We shall show later (§5) that the sex ratio version of the TWH holds under quite general conditions on  $f_{\sigma}$  and  $f_{\varphi}$ . These conditions are satisfied by the linear fitness functions studied in this section. Still, owing to the simplicity of the present fitness functions, it will nonetheless be instructive to demonstrate here that the sex ratio version of the TWH holds for them.

It is proved in the electronic supplementary material, appendix S1 that a mother of condition  $c_m$  will have an optimal sex ratio that is male-biased if and only if she would prefer, in fitness terms, to have two male offspring than two female offspring (again, we stress that this comparison of the fitness sums of an all-male brood and an all-female brood is a diagnostic test for whether the mother's optimal sex ratio is *at all* male- or female-biased, and does not imply that her optimal sex ratio is necessarily one or zero). Note that she always apportions investment equitably among a same-sex brood, owing to the concavity of fitness with respect to investment for both male and female offspring (a result of condition being concave in investment and this concavity being maintained under the linear transformation to fitness). So the condition for a male-biased sex ratio amounts to the following inequality:

$$\begin{aligned} 2f_{\sigma}(c_m, I(c_m)/2) &> 2f_{\varphi}(c_m, I(c_m)/2) \\ \Leftrightarrow 2\lambda c(c_m, I(c_m)/2) &> 2[c(c_m, I(c_m)/2) + k] \\ \Leftrightarrow c(c_m, I(c_m)/2) &> k/(\lambda - 1). \end{aligned}$$

But as  $(\partial/\partial c_m)c(c_m, I(c_m)/2) = \partial_1 c(c_m, I(c_m)/2) + \partial_2 c(c_m, I(c_m)/2)I'(c_m)/2 > 0$ , and assuming the existence of some  $c_m^* \in (0, 1)$  such that  $c(c_m^*, I(c_m^*)/2) = k/(\lambda - 1) =: c^*$ , this implies that mothers with condition above  $c_m^*$  have a male-biased optimal sex ratio, and mothers with condition below  $c_m^*$  have a female-biased optimal sex ratio. The sex ratio version of the TWH therefore holds in this case, while the investment version does not.

Finally, it is interesting that mothers of condition  $c_m^*$ , though indifferent between having an all-male brood and an all-female brood, are not indifferent between those choices and having a mixed brood. To see this, note that each male in an all-male brood of such a mother has the same fitness as each female in an all-female brood, and each receives the same investment  $I(c_m^*)/2$ . Thus, a mother with a mixed brood could achieve the same fitness sum as with a same-sex brood if she invests  $I(c_m^*)/2$  in each offspring, male and female. But as the argument above demonstrates, this is not

the optimal investment decision for the mother: she should always invest more in the male offspring than in the female offspring of a mixed brood. So her fitness sum from having a mixed brood and apportioning investment optimally is strictly higher than her fitness sum from having a same-sex brood and apportioning investment optimally; and as a result, by continuity, mothers whose condition is sufficiently close to  $c_m^*$ , though the bias in their sex ratio is determined by whether their condition is higher or lower than this value, would nonetheless have optimal sex ratios that are mixed (i.e. not zero or one). The intuition for this result is that, with a mixed brood, the mother can divert investment away from the female offspring, for whom it is relatively unproductive (in fitness terms), towards the male offspring, for whom it is relatively productive. This choice is not available to the mother of a same-sex brood.

This explanation complements other reasons for why selection favours individual sex ratios that are not zero or one. For example, in situations with local within-sex competition for mates, selection can favour extreme sex ratios [45], but it will also not favour those that are zero or one. It also suggests an additional benefit of precise control of brood composition (i.e. 'homeostatic', as opposed to probabilistic, sex ratios) over the usually-cited lower-variance benefit [46,47], which is known to be small [47,48].

## 4. Broad sufficient conditions for the sex ratio version to hold

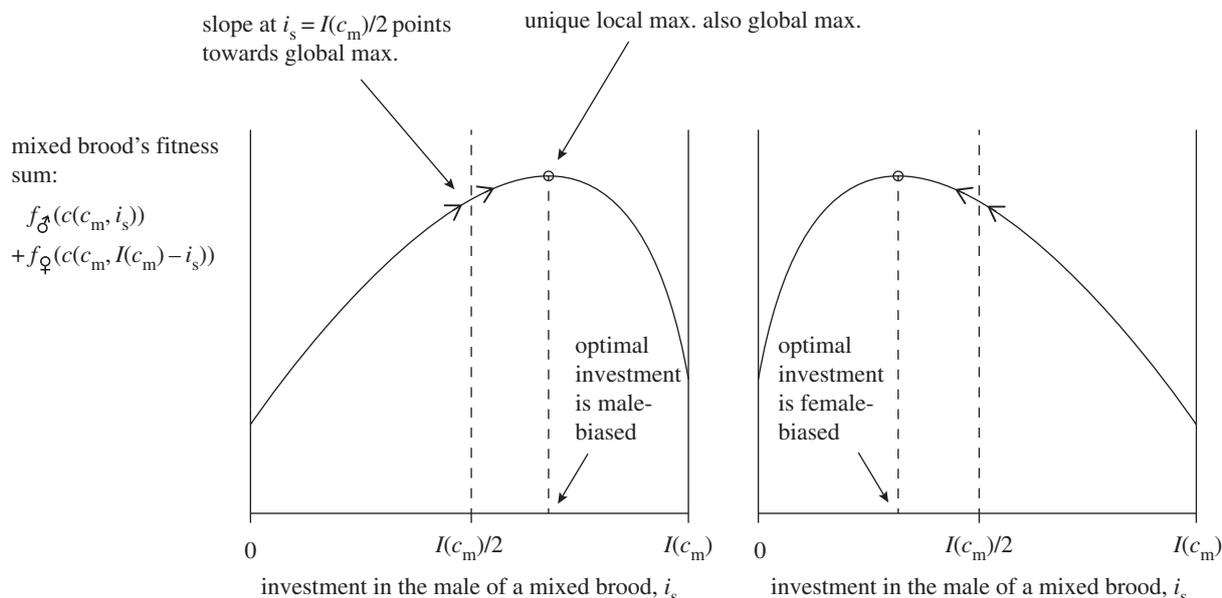
Assume that the optimal investment decision for the mother of a same-sex brood, no matter her condition, is always to apportion investment equitably between her two offspring. This holds if the functions  $f_{\sigma}(c(c_m, i))$  and  $f_{\varphi}(c(c_m, i))$  are both concave in investment  $i$ —loosely speaking, given that  $c(c_m, i)$  is already concave in  $i$ , this means that either  $f_{\sigma}$  and  $f_{\varphi}$  are themselves (weakly) concave, or that any convexities in them are not sufficient to overturn the concavity of  $c(c_m, i)$  in  $i$ . This is satisfied, for example, by the linear fitness functions assumed in §3, by those displayed in figure 4, and by those displayed in figure 3 provided the convex fitness function  $f_{\sigma}(c)$  is not too convex.<sup>4</sup>

With this assumption, a simple and natural set of 'single-crossing' conditions under which the sex ratio version of the TWH always holds is as follows.

- (SC1) The condition-fitness profiles for males and females cross only once (at  $c^*$ ), with the male profile starting below the female profile (in poor adult condition, males are less fit than females:  $c < c^* \Rightarrow f_{\sigma}(c) < f_{\varphi}(c)$ ) and ending above it (in good adult condition, males are fitter than females:  $c > c^* \Rightarrow f_{\sigma}(c) > f_{\varphi}(c)$ );
- (SC2) There exists some maternal condition  $c_m^*$  such that  $c^* = c(c_m^*, I(c_m^*)/2)$ .

If  $f_{\sigma}(c(c_m, i))$  and  $f_{\varphi}(c(c_m, i))$  are concave in  $i$ , and if (SC1) and (SC2) hold, then mothers in condition  $c_m > c_m^*$  have optimal sex ratios that are male-biased, while those in condition  $c_m < c_m^*$  have optimal sex ratios that are female-biased.

To prove this, first consider the case of a female in condition  $c_m > c_m^*$ . A sufficient condition for her optimal sex ratio to be male-biased is that she would achieve higher fitness with an all-male brood than with an all-female



**Figure 2.** If the fitness sum of a mixed brood is ‘single-humped’ with respect to investment in the male of the brood, then the slope of the function at the point of equitable investment points in the direction of the optimal bias in investment, significantly simplifying detection of the latter in our model.

brood (as shown in the electronic supplementary material, appendix S1). As her optimal investment decision in either case is equitable apportionment among the brood, this condition becomes:

$$2f_{\sigma}(c(c_m, I(c_m)/2)) > 2f_{\varphi}(c(c_m, I(c_m)/2)).$$

But as  $(\partial/\partial c_m)c(c_m, I(c_m)/2) = \partial_1 c(c_m, I(c_m)/2) + \partial_2 c(c_m, I(c_m)/2)I'(c_m)/2 > 0$ , we have that  $c_m > c_m^* \Rightarrow c(c_m, I(c_m)/2) > c(c_m^*, I(c_m^*)/2) = c^*$ , and so, from the single-crossing condition (SC1),  $f_{\sigma}(c(c_m, I(c_m)/2)) > f_{\varphi}(c(c_m, I(c_m)/2))$ .

A female in condition  $c_m < c_m^*$  would, by a similar argument, achieve higher fitness with an all-female brood than with an all-male brood, and so her optimal sex ratio would be female-biased.

The case where at least one of  $f_{\sigma}(c(c_m, i))$  or  $f_{\varphi}(c(c_m, i))$  is not everywhere concave in  $i$  is more complicated, and in general, the single-crossing conditions do not guarantee that the sex ratio version holds (a graphical counterexample is given in the electronic supplementary material, appendix S2).

## 5. Narrow sufficient conditions for the investment version to hold

We have seen in §3 that, for a simple set of fitness functions that satisfy the Trivers–Willard assumptions, the investment version of the hypothesis does not hold. In this section, we derive conditions under which the investment version of the TWH does hold. They are significantly narrower than the basic set of assumptions stated by Trivers & Willard [1].

To simplify the analysis, we assume that to determine the investment bias (male or female) in a mixed brood of a mother in condition  $c_m$ , it suffices to check in which direction natural selection ‘points’ from an initial state of equitable investment in the male and female offspring (i.e. does it point towards increased, and thus biased, investment in the male or in the female?). If the ancestral state is equitable apportionment of investment within mixed broods, and if

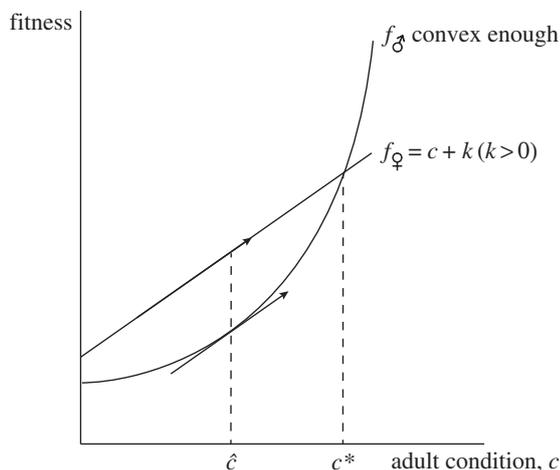
maternal behavioural changes are gradual, then this is the investment bias that natural selection would lead to.

Mathematically, this amounts to checking the slope of the mixed brood’s fitness sum  $f_{\sigma}(c(c_m, i_s)) + f_{\varphi}(c(c_m, I(c_m) - i_s))$  with respect to investment  $i_s$  in the male offspring, evaluated at the point  $i_s = I(c_m)/2$ . If the brood’s fitness sum is of the ‘single-humped’ shape displayed in figure 2, i.e. with one local maximum that is also the global maximum, and with no local minima, then this slope indeed points towards the global maximum.

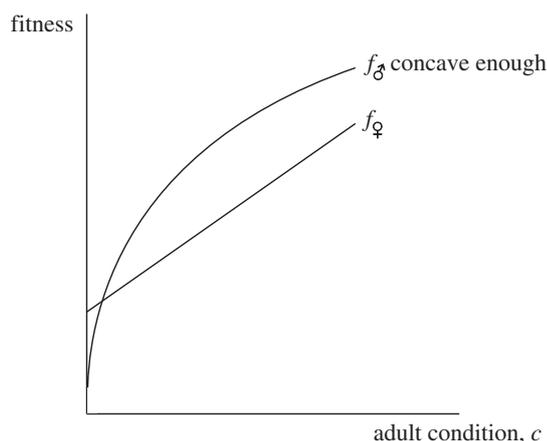
The investment version of the TWH then simplifies to the following: the partial derivative of the fitness sum of a mixed brood, taken with respect to investment in the male offspring, and evaluated at the point of equitable investment  $I(c_m)/2$ , is negative for maternal conditions below a certain threshold, and positive for maternal conditions above that threshold. This partial derivative, in general form, is given by:

$$\begin{aligned} & \left. \frac{\partial}{\partial i_s} [f_{\sigma}(c(c_m, i_s)) + f_{\varphi}(c(c_m, I(c_m) - i_s))] \right|_{i_s=I(c_m)/2} \\ &= f'_{\sigma}(c(c_m, I(c_m)/2))\partial_2 c(c_m, I(c_m)/2) \\ & \quad - f'_{\varphi}(c(c_m, I(c_m)/2))\partial_2 c(c_m, I(c_m)/2) \\ &= [f'_{\sigma}(c(c_m, I(c_m)/2)) - f'_{\varphi}(c(c_m, I(c_m)/2))] \partial_2 c(c_m, I(c_m)/2). \end{aligned} \quad (5.1)$$

As  $\partial_2 c(c_m, I(c_m)/2) > 0$ , the sign of this expression is the same as the sign of the term in square brackets in the last line. The investment version of the TWH thus further simplifies to the following:  $f'_{\sigma}(c(c_m, I(c_m)/2)) < f'_{\varphi}(c(c_m, I(c_m)/2))$  for all  $c_m$  below some threshold, and  $f'_{\sigma}(c(c_m, I(c_m)/2)) > f'_{\varphi}(c(c_m, I(c_m)/2))$  for all  $c_m$  above that threshold. In other words, male fitness responds less to increases in adult condition than does female fitness below some threshold adult condition, and male fitness responds more to changes in adult condition than does female fitness above that threshold condition. This condition is rather specific, and will hold for only a subset of the fitness functions that are consistent with Trivers and Willard’s original assumptions. Indeed, under a strict reading of Trivers and Willard’s ‘Assumption 3’ (quoted in §2), it appears that the condition



**Figure 3.** Fitness functions for which the investment version of the Trivers–Willard hypothesis does hold. If the male fitness function is sufficiently convex, then the very small slope at low conditions makes investment in males in low condition unattractive relative to investment in females, and so mothers in low condition should bias investment towards females. Conversely, investment in males in good condition is very attractive (their fitness function is very steep), and so mothers in good condition should bias their investment towards male offspring.



**Figure 4.** Fitness functions for which a situation opposite that predicted by the investment version of the Trivers–Willard hypothesis is expected to result. Because the slope of the male fitness function starts out very steep, investment in males in poor condition yields high fitness returns and so mothers in poor condition would be expected to bias investment towards the male of a mixed brood. Conversely, mothers in good condition are expected to bias investment towards the female of a mixed brood, because the eventual flatness of the male fitness function makes investment in males in good condition relatively unproductive.

(5.1) we have derived for the investment version to hold is in fact disallowed! Nevertheless, there are some natural instances where we might expect them to be fulfilled, as the following example illustrates.

We keep the female fitness function from before,  $f_{\text{♀}}(c) = c + k$ ,  $k > 0$ , and specify a convex male fitness function  $f_{\text{♂}}(c)$  such that the slope of  $f_{\text{♂}}(c)$  is less than one (the slope of  $f_{\text{♀}}(c)$ ) for all adult conditions below a certain (realized) threshold  $\hat{c}$  (we keep the notation  $c^*$  for the point at which  $f_{\text{♂}}$  and  $f_{\text{♀}}$  cross), and greater than one for all adult conditions above that threshold (figure 3). That is, we require the male fitness function to be *sufficiently convex*. Under this scenario,

it is clear that the condition (5.1) is satisfied, and we expect the investment version of the TWH to hold in this case.

If we instead specify that the male fitness function be sufficiently concave, while keeping the female fitness function linear (figure 4), then a situation results that is directly opposed to that predicted by the investment version of the TWH. Now, mothers in poor condition are expected to bias investment towards the male of a mixed brood, while mothers in good condition are expected to bias investment towards the female of a mixed brood. Nonetheless, the sex ratio version of the TWH is expected still to hold in this case, because the single crossing conditions (SC1) and (SC2) hold.

## 6. Discussion

We have demonstrated, in the context of a simple unified model, that the conditions under which the sex ratio version of the TWH holds are significantly broader than those under which the investment version is expected to hold. This helps to make sense of the observation that empirical tests of the former tend to return positive results, while empirical tests of the latter return more mixed results [20].

The difference between the two versions of the hypothesis can be understood in terms of the economic difference between absolute and marginal value. Whether a mother in a certain condition should bias her sex ratio towards male offspring amounts to asking whether the male fitness function *lies above* the female fitness function for the relevant condition, i.e. whether there exists an absolute difference in fitness between offspring of the two sexes. On the other hand, for a mother with a mixed brood, whether she should bias parental investment towards the male offspring amounts to asking whether the male fitness function is *steeper* than the female fitness function for the relevant condition, i.e. whether improvement of the male's condition yields a greater marginal fitness return. This insight is clearly captured by Keller *et al.* [20, pp. 356–357], who find no evidence for an investment version Trivers–Willard effect in the USA:

[Resource allocation biasing] should be biased toward the offspring that most improves the parents' fitness per unit invested. However, this bias has no necessary link to the fitness value of the offspring. Rather, it is predicted by cues related to the marginal gains of allocating each particular resource to the sex in question. The assumptions of the TWH (a higher correlation between condition and [reproductive success] among males than females, and a correlation between the parent's and offspring's conditions) are unrelated to these marginal benefits, but they *are* related to the fitness value of offspring.

Loosely, then, the sex ratio version of the TWH is expected to hold when the male and female fitness functions are such that males in poor condition have relatively low fitness (the male fitness function starts out below the female fitness function), males in good condition have relatively high fitness (the male fitness function ends up above the female fitness function) and the fitness functions cross once (e.g. figures 1, 3 and 4). This describes the classic picture of male and female fitness functions in polygynous species, and so we should expect the sex ratio version of the TWH to hold very generally.

The investment version, on the other hand, is expected to hold only if the male fitness function is much flatter than the female fitness function for poor conditions, but much steeper

for good conditions (e.g. figure 3). This condition is much more restrictive than that under which the sex ratio version holds, and so we should expect the investment version of the TWH to hold less generally than the sex ratio version.

Our model is simply intended to capture the logic of why the sex ratio and investment versions of the TWH differ. We discuss some of the model's assumptions, and the possibility of their relaxation, below.

First, our analysis optimizes over phenotypes, and therefore ignores both their genetic basis and the evolutionary dynamics that might lead to their optimization. In this sense, it is a heuristic analysis. Studying phenotypes in a static, non-genetic setting was vital for simplicity of the model (for a defence of this style of modelling, see Parker & Maynard Smith [49]). Nonetheless, it remains possible that the optimal behaviours derived here are not evolvable in a simple population genetic framework.

Second, the assumption of a single fixed brood of constant size two is, of course, not realistic in general, having been made for the sake of tractability. One direction in which the model could be made more realistic would be to allow brood size to vary, and to undergo selection in conjunction with sex ratio and the pattern of parental investment. While it is not clear whether this could be done in a way that still permitted tractable analysis, it could alter the interpretation of our results if there were significant interactions between optimal brood size and the optima of our variables of interest, sex ratio and the pattern of parental investment ([48], [50,59]).

In this vein, another way in which the model could be generalized would be to consider sex ratio and investment patterns across sequential broods. Consider the simplest case that of two offspring raised one after the other. The optimization problem faced by the mother is then slightly different to that she faces in our single-brood model. For example, she might be able to alter her sex ratio for the second offspring conditional on the sex of her first offspring; this option is not available to the mother of a single brood. In this case, the results of the sequential model are very similar to those of the single-period model we have analysed (electronic supplementary material, appendix S3). If, on the other hand, her sex ratio cannot change between periods in the sequential model, then a mother whose sex ratio is not zero or one invests in her first offspring without knowing the sex of her second offspring. This case is significantly more complicated, but a tentative analysis (electronic supplementary material, appendix S3) suggests that its results would usually not differ significantly from those of our single-period model.

Third, our model is, in the language of economic theory, a 'partial equilibrium' analysis: the decision problem faced by one mother is independent of the decisions made by other mothers. In our model, for example, the exogenous parameter  $k$ , mediating the height of the female fitness function, can be adjusted so that, given the population distribution of maternal condition, the average reproductive success of sons and daughters is equal. In reality, the mating success of a male in good condition depends on the number of other males in good condition, and this in turn depends on the sex allocation decisions of other mothers in the population. A full general equilibrium model would specify a distribution of maternal conditions, and sex ratio and investment strategies for mothers in each condition,

such that: (i) each mother's strategies are optimal, given the strategies of other mothers, (ii) the average reproductive success of all sons would match that of all daughters, and possibly (iii) the distribution of adult condition of the daughters matches that of the mothers. We do not believe that the technical tools of sex allocation theory are yet advanced enough to approach this general equilibrium problem, but it represents an important avenue of future research.<sup>5</sup>

Fourth, our model treats maternal and offspring condition as continuous variables, where in reality the important factor for reproductive success may be discrete, like rank within a social hierarchy. The strategic concerns raised in the previous paragraph apply with even more force here: whether a mother's investment in her son causes him to jump up one place in the male condition ranking clearly depends on the decision of the mother of the son presently 'occupying' that place. A more subtle complication is that the discretization of the fitness functions' domain (here, within-sex rank) makes analysis of the effects of continuous variables (like investment) on fitness significantly less tractable.

A related, but more fundamental problem with the approach adopted here is that 'condition' is treated as an abstract, and therefore unobserved, quantitative variable. But many of the contingencies that we have shown to be of theoretical importance involve the shape of reproductive value functions with respect to condition; as condition is not empirically observed, neither can the shape of these functions be. The criteria that we have derived for when the sex ratio and investment versions of the TWH hold are therefore not directly measurable.

Getting around this problem would require substitution of 'condition' for an imperfect but observable proxy variable, such as weight. That is, it would require measurement of the fitness functions with respect to weight, and how weight responds to parental condition and investment. This would detract significantly from the generality of the theory, and would be accurate only insofar as weight correlates closely with true condition. It is possible that a composite proxy involving many variables (such as weight, nutritional status, sexual ornaments, etc.) might improve this correlation.

This problem applies not only to our model, but to all theoretical models of the TWH that treat 'condition' as an intermediate variable [36]. It appears to us to be the chief obstacle to generating testable predictions from such models, and its solution is therefore a crucial line of future research.

**Authors' contributions.** C.V. conceived the study and carried out modelling work. C.V., D.H. and M.A.N. analysed results and wrote the manuscript.

**Competing interests.** We have no competing interests.

**Funding.** We received no funding for this study.

**Acknowledgement.** We are grateful to Kirill Borusyak, Philip Downing, Jean-Michel Gaillard, Robert Trivers, Stuart West, and the referees for helpful comments.

## Endnotes

<sup>1</sup>Insignificant results can be explained either by inadequate sample size or constraints imposed by genetic sex determination. Significant results in the opposite direction can be explained by demographic factors [22,23].

<sup>2</sup>Note that all results in this paper hold in the simpler cases where (i) the improved condition of offspring of high-condition mothers

derives only from those mothers' ability to invest more in their offspring ( $I'(c_m) > 0$ ,  $(\partial/\partial c_m)c(c_m, i) = 0$  and  $(\partial/\partial i)c(c_m, i) > 0$ ), and (ii) there is no difference in investment ability between high and low quality mothers ( $I'(c_m) = 0$  but  $(\partial/\partial c_m)c(c_m, i) > 0$ ).

<sup>3</sup>This result, and the reason behind it, extend to any functions  $f_{\sigma}$  and  $f_{\varphi}$  that satisfy  $f'_{\varphi}(x) < f'_{\sigma}(y)$  for any  $x$  and  $y$ , since then the general first-order condition for  $i_s^*$  yields

$$f'_{\sigma}(c(c_m, i_s^*))\partial_2 c(c_m, i_s^*) - f'_{\varphi}(c(c_m, I(c_m) - i_s^*))\partial_2 c(c_m, I(c_m) - i_s^*) = 0$$

$$\Rightarrow \frac{\partial_2 c(c_m, i_s^*)}{\partial_2 c(c_m, I(c_m) - i_s^*)} = \frac{f'_{\sigma}(c(c_m, I(c_m) - i_s^*))}{f'_{\varphi}(c(c_m, i_s^*))} < 1 \Rightarrow i_s^* > I(c_m) - i_s^*,$$

where the last step follows from the concavity of  $c$  with respect to its second argument.

<sup>4</sup>More precisely, in this case of convex  $f_{\sigma}(c)$ , we require  $(\partial^2/\partial i^2)f_{\sigma}(c(c_m, i)) < 0$ , which translates to the condition  $|\partial_{22}c(c_m, i)/\partial_2 c(c_m, i)| > |f''_{\sigma}(c(c_m, i))/f'_{\sigma}(c(c_m, i))|$ , i.e., that  $c$  is relatively more concave in  $i$  than  $f_{\sigma}$  is convex in  $c$ .

<sup>5</sup>On the strategic considerations, however, we expect that, in equilibrium, a certain monotonicity condition should hold: in particular, within-sex offspring condition rank should covary perfectly with maternal condition rank. If this were not satisfied, we would have the odd scenario that, in equilibrium, a mother in better condition than another neglects her son (say) sufficiently much relative to the other mother that her son's post-investment condition is lower than that of the other mother's son. So, if this monotonicity condition were to hold, we should not expect strategic concerns to alter the structure of the problem, at least in terms of the rank of offspring condition.

## References

- Trivers RL, Willard DE. 1973 Natural selection of parental ability to vary the sex ratio of offspring. *Science* **179**, 90–92. (doi:10.1126/science.179.4068.90)
- Clutton-Brock TH, Albon SD, Guinness FE. 1984 Maternal dominance, breeding success and birth sex ratios in red deer. *Nature* **308**, 358–360. (doi:10.1038/308358a0)
- Clutton-Brock TH, Albon SD, Guinness FE. 1986 Great expectations: dominance, breeding success and offspring sex ratios in red deer. *Anim. Behav.* **34**, 460–471. (doi:10.1016/S0003-3472(86)80115-4)
- Cameron EZ, Linklater WL. 2007 Extreme sex ratio variation in relation to change in condition around conception. *Biol. Lett.* **3**, 395–397. (doi:10.1098/rsbl.2007.0089)
- Charnov EL, Los-den Hartogh RL, Jones WT, van den Assem J. 1981 Sex ratio evolution in a variable environment. *Nature* **289**, 27. (doi:10.1038/289027a0)
- Mealey L, Mackey W. 1990 Variation in offspring sex ratio in women of differing social status. *Ethol. Sociobiol.* **11**, 83–95. (doi:10.1016/0162-3095(90)90030-A)
- Cameron EZ, Dalerum F. 2009 A Trivers-Willard effect in contemporary humans: male biased sex ratios among billionaires. *PLoS ONE* **4**, e4195. (doi:10.1371/journal.pone.0004195)
- Fujita M, Roth E, Lo YJ, Hurst C, Vollner J, Kendell A. 2012 In poor families, mothers' milk is richer for daughters than sons: a test of Trivers-Willard hypothesis in agropastoral settlements in Northern Kenya. *Am. J. Phys. Anthropol.* **149**, 52–59. (doi:10.1002/ajpa.22092)
- Cronk L. 1989 Low socioeconomic status and female-biased parental investment: the Mukogodo example. *Am. Anthropol.* **91**, 414–429. (doi:10.1525/aa.1989.91.2.02a00090)
- Gaulin SJC, Robbins CJ. 1991 Trivers-Willard effect in contemporary North American society. *Am. J. Phys. Anthropol.* **85**, 61–69. (doi:10.1002/ajpa.1330850108)
- Berezkei T, Dunbar RIM. 1997 Female-biased reproductive strategies in a Hungarian Gypsy population. *Proc. R. Soc. Lond. B* **264**, 17–22. (doi:10.1098/rspb.1997.0003)
- Hewison AJM, Gaillard JM. 1999 Successful sons or advantaged daughters? The Trivers-Willard model and sex-biased maternal investment in ungulates. *Trends Ecol. Evol.* **14**, 229–234. (doi:10.1016/S0169-5347(99)01592-X)
- Cameron EZ, Linklater WL. 2000 Individual mares bias investment in sons and daughters in relation to their condition. *Anim. Behav.* **60**, 359–367. (doi:10.1006/anbe.2000.1480)
- Brown GR. 2001 Sex-biased investment in nonhuman primates: can Trivers & Willard's theory be tested? *Anim. Behav.* **61**, 683–694. (doi:10.1006/anbe.2000.1659)
- Quinlan RJ, Quinlan MB, Flinn MV. 2003 Parental investment and age at weaning in a Caribbean village. *Evol. Hum. Behav.* **24**, 1–16. (doi:10.1016/S1090-5138(02)00104-6)
- Hinde K. 2007 First-time macaque mothers bias milk composition in favor of sons. *Curr. Biol.* **17**, R958–R959. (doi:10.1016/j.cub.2007.09.029)
- Hinde K. 2009 Richer milk for sons but more milk for daughters: sex-biased investment during lactation varies with maternal life history in rhesus macaques. *Am. J. Hum. Biol.* **21**, 512–519. (doi:10.1002/ajhb.20917)
- Wander K, Mattison SM. 2013 The evolutionary ecology of early weaning in Kilimanjaro, Tanzania. *Proc. R. Soc. B* **280**, 20131359. (doi:10.1098/rspb.2013.1359)
- Cameron EZ, Linklater WL. 2002 Sex bias in studies of sex bias: the value of daughters to mothers in poor condition. *Anim. Behav.* **63**, F5–F8. (doi:10.1006/anbe.2001.1902)
- Keller MC, Nesse RM, Hofferth S. 2001 The Trivers-Willard hypothesis of parental investment: no effect in the contemporary United States. *Evol. Hum. Behav.* **22**, 343–360. (doi:10.1016/S1090-5138(01)00075-7)
- Sheldon BC, West SA. 2004 Maternal dominance, maternal condition, and offspring sex ratio in ungulate mammals. *Am. Nat.* **163**, 40–54. (doi:10.1086/381003)
- Leimar O. 1996 Life-history analysis of the Trivers and Willard sex-ratio problem. *Behav. Ecol.* **7**, 316–325. (doi:10.1093/beheco/7.3.316)
- Schindler S, Gaillard JM, Gruning A, Neuhaus P, Traill LW, Tuljapurkar S, Coulson T. 2015 Sex-specific demography and generalization of the Trivers-Willard theory. *Nature* **526**, 249–252. (doi:10.1038/nature14968)
- Edwards AWF. 1958 An analysis of Geissler's data on the human sex ratio. *Ann. Hum. Genet.* **23**, 6–15. (doi:10.1111/j.1469-1809.1958.tb01437.x)
- Almond D, Edlund L. 2007 Trivers-Willard at birth and one year: evidence from US natality data 1983–2001. *Proc. R. Soc. B* **274**, 2491–2496. (doi:10.1098/rspb.2007.0524)
- Carranza J. 2002 What did Trivers and Willard really predict? *Anim. Behav.* **63**, F1–F3. (doi:10.1006/anbe.2001.1901)
- Frank SA. 1987 Individual and population sex allocation patterns. *Theor. Popul. Biol.* **31**, 47–74. (doi:10.1016/0040-5809(87)90022-0)
- Düsing C. 1883 Die Faktoren welche die Sexualität entscheiden. *Jenaische Z. Nat.wiss.* **16**, 428–464.
- Fisher RA. 1930 *The genetical theory of natural selection*. Oxford, UK: Clarendon Press.
- Maynard Smith J. 1980 A new theory of sexual investment. *Behav. Ecol. Sociobiol.* **7**, 247–251. (doi:10.1007/BF00299371)
- Lessells CM. 1998 A theoretical framework for sex-biased parental care. *Anim. Behav.* **56**, 395–407. (doi:10.1006/anbe.1998.0764)
- Godfray HCJ. 1986 Models for clutch size and sex ratio with sibling interaction. *Theor. Popul. Biol.* **30**, 215–231. (doi:10.1016/0040-5809(86)90034-1)
- Nagelkerke CJ. 1994 Simultaneous optimization of egg distribution and sex allocation in a patch-structured population. *Am. Nat.* **144**, 262–284. (doi:10.1086/285674)
- Greeff JM. 1997 Offspring allocation in externally ovipositing fig wasps with varying clutch size and sex ratio. *Behav. Ecol.* **8**, 500–505. (doi:10.1093/beheco/8.5.500)
- West SA, Flanagan KE, Godfray HCJ. 1999 Sex allocation and clutch size in parasitoid wasps that produce single-sex broods. *Anim. Behav.* **57**, 265–275. (doi:10.1006/anbe.1998.0958)
- Clancey E, Byers JA. 2014 The definition and measurement of individual condition in evolutionary studies. *Ethology* **120**, 845–854. (doi:10.1111/eth.12272)

37. Hausfater G, Altmann J, Altmann S. 1982 Long-term consistency of dominance relations among female baboons (*Papio cynocephalus*). *Science* **217**, 752–755. (doi:10.1126/science.217.4561.752)
38. East ML, Höner OP, Wachter B, Wilhelm K, Burke T, Hofer H. 2009 Maternal effects on offspring social status in spotted hyenas. *Behav. Ecol.* **20**, 478–483. (doi:10.1093/beheco/arp020)
39. Moss R, Rothery P, Trenholm IB. 1985 The inheritance of social dominance rank in red grouse (*Lagopus lagopus scoticus*). *Aggress. Behav.* **11**, 253–259. (doi:10.1002/1098-2337(1985)11:3<253::AID-AB2480110308>3.0.CO;2-Z)
40. Pen I, Weissing FJ. 2001 Sexual selection and the sex ratio: an ESS analysis. *Selection* **1**, 111–122. (doi:10.1556/Select.1.2000.1-3.11)
41. Ellegren H, Gustafsson L, Sheldon BC. 1996 Sex ratio adjustment in relation to paternal attractiveness in a wild bird population. *Proc. Natl Acad. Sci. USA* **93**, 11 723–11 728. (doi:10.1073/pnas.93.21.11723)
42. Fawcett TW, Kuijper B, Pen I, Weissing FJ. 2007 Should attractive males have more sons? *Behav. Ecol.* **18**, 71–80. (doi:10.1093/beheco/arl052)
43. Kuijper B, Pen I, Weissing FJ. 2012 A guide to sexual selection theory. *Annu. Rev. Ecol. Evol. Syst.* **43**, 287–311. (doi:10.1146/annurev-ecolsys-110411-160245)
44. Haig D. 1990 Brood reduction and optimal parental investment when offspring differ in quality. *Am. Nat.* **136**, 550–556. (doi:10.1086/285113)
45. Hamilton WD. 1967 Extraordinary sex ratios. *Science* **156**, 477–488. (doi:10.1126/science.156.3774.477)
46. Verner J. 1965 Selection for sex ratio. *Am. Nat.* **99**, 419–421. (doi:10.1086/282384)
47. Taylor PD, Sauer A. 1980 The selective advantage of sex-ratio homeostasis. *Am. Nat.* **116**, 305–310. (doi:10.1086/283629)
48. Williams GC. 1979 The question of adaptive sex ratio in outcrossed vertebrates. *Proc. R. Soc. Lond. B* **205**, 567–580. (doi:10.1098/rspb.1979.0085)
49. Parker GA, Maynard Smith J. 1990 Optimality theory in evolutionary biology. *Nature* **348**, 27–33. (doi:10.1038/348027a0)
50. Frank SA. 1998 *Foundations of social evolution*. Princeton, NJ: Princeton University Press.