Author(s): Fromhage, Lutz; Henshaw, Jonathan M.

Title: The balance model of honest sexual signaling

Year: 2022

Version: Published version

Please cite the original version:

Fromhage, L., & Henshaw, J. M. (2022). The balance model of honest sexual signaling. Evolution,
76(3), 445-454. https://doi.org/10.1111/evo.14436

The balance model of honest sexual signaling

Lutz Fromhage[1,2] (iD) and Jonathan M. Henshaw[3] (iD)

[1]*Department of Biological and Environmental Science, University of Jyvaskyla, Jyvaskyla 40014, Finland*

[2]*E-mail: lutz.fromhage@jyu.fi*

[3]*Institute of Biology I (Zoology), University of Freiburg, Freiburg 79104, Germany*

Costly signaling theory is based on the idea that individuals may signal their quality to potential mates and that the signal's costliness plays a crucial role in maintaining information content ("honesty") over evolutionary time. Although costly signals have traditionally been described as "handicaps," here we present mathematical results that motivate an alternative interpretation. We show that under broad conditions, the multiplicative nature of fitness selects for roughly balanced investments in mating success and viability, thereby generating a positive correlation between signal size and quality. This balancing tendency occurs because selection for increased investment in a fitness component diminishes with the absolute level of investment in that component, such that excessively biased investments are penalized. The resulting interpretation of costly signals as balanced (albeit not necessarily equal) investments may be a widely applicable alternative to the traditional "handicap" metaphor, which has been criticized for its non-Darwinian connotation of selection for "waste" rather than efficiency. We predict that accelerating returns on viability are necessary to undermine honesty. This prediction depends crucially on the assumption that mating success and viability contribute multiplicatively (rather than additively) to an individual's fitness.

In many animal species, males (or, more rarely, females) exhibit conspicuous secondary sexual traits that are targets of mate choice by the opposite sex, potentially because they provide reliable information about an individual's condition, health, and/or social status (Andersson 1994; Dougherty 2021). This raises the question: what prevents poor-quality individuals from "cheating", that is, expressing traits that give an exaggerated impression of their actual quality? A popular answer to this question is Amotz Zahavi's handicap hypothesis (Zahavi 1975; reviewed in Penn and Számadó 2020), which states that signal traits impose survival costs that can more easily be borne by high-quality individuals. The handicap hypothesis was initially met with skepticism (e.g., Kirkpatrick 1986), but became widely accepted after Grafen (1990) presented a theoretical model showing that honest signaling can be stable when individuals strategically choose their signal strength conditional on their quality. Although some researchers took Grafen's model as setting a new benchmark for interpreting the handicap hypothesis, Zahavi's writings remained

a source of confusion. This is because Zahavi applied his handicap terminology to several distinct ideas, some of which are incompatible with Darwinian logic (Penn and Számadó 2020). In particular, Zahavi claimed that signals evolve through a special kind of selection which—in contrast to natural selection—favors waste and inefficiency (Zahavi 1981, 1987; Zahavi and Zahavi 1997). This claim reflects the obsolete view that natural selection maximizes individual survival rather than gene propagation, such that a trait that reduces survival counts as wasteful regardless of its effect on reproduction. Although Grafen's (1990) model was perfectly Darwinian, he nevertheless embraced the Zahavian rhetoric of wastefulness, stating, for example (p. 532): "The evolutionary stability of persuasive signaling necessitates honesty, which necessitates waste." Grafen (1990) also derived a more specific result, which states that, for honesty to be stable, the marginal cost of advertising should be greater for worse males (the "decreasing-marginal-cost criterion"; Getty 2006). Accordingly, borrowing an example from Getty (2006, p. 84),

"a high-quality peacock with 100 eye spots on its tail should be able to add one more spot at less viability cost than could a poor-quality peacock with 100 eye spots." Yet, as Getty (2006) pointed out, the decreasing-marginal-cost criterion rests on the implicit assumption that the costs and benefits of signaling combine in an additive manner—which cannot be true if benefits arise in terms of mating success and costs arise in terms of viability. Without the unrealistic assumption of additivity, the decreasing-marginal-cost criterion boils down to the more general criterion that the marginal net fitness benefit of increased signal strength should increase with quality (Getty 2006): that is, a high-quality peacock with 100 eye spots should gain more fitness from adding one more spot than could a poor-quality peacock with 100 eye spots. Unfortunately, this criterion is so abstract that it is unclear how easily it might be satisfied (Getty 2006). Because there is no logical necessity for net fitness benefits to increase in this way, one may get the impression that honest signaling only works if the quality-dependent fitness functions happen to have very specific shapes by coincidence (e.g., a product of exponential functions, as assumed by Iwasa et al. 1991; Johnstone et al. 2009).

By contrast, here we argue that there is a powerful reason why fitness functions should often be of the right kind to support honest sexual signaling. We locate this reason in the multiplicative nature of the mating versus viability trade-off that is commonly supposed to underlie costly sexual signaling. To develop our argument, let us briefly step back to compare different kinds of trade-offs in a more general setting.

Additive and Multiplicative Trade-Offs

A trade-off describes a situation in which two or more desirable (i.e., fitness-enhancing) quantities cannot be maximized simultaneously, requiring a decision that balances conflicting needs. We will focus in particular on trade-offs where fitness can be expressed either as the sum of fitness components ("additive" trade-offs) or as the product of fitness components ("multiplicative" trade-offs; also see Houston et al. 2003). One way to formalize the idea that fitness components $x_1(u_1)$, $x_2(u_2)$, $\ldots$, $x_n(u_n)$ are subject to a trade-off is to postulate a constraint $R = u_1 + u_2 + \cdots + u_n$ that must be met by the allocations u_1, u_2, $\ldots$, u_n to these components (with $u_i \geq 0$). The simplest interpretation of this is that R represents the total amount of some limiting resource (e.g., energy: Somjee 2021) accumulated by the organism before the allocation decision. More generally, R can be envisaged as an abstract measure of "quality" that constrains the multivariate phenotypic space that an individual can access via its developmental and behavioral decisions. Hence, R may also capture an organism's capacity to acquire resources and to use them efficiently to express adaptive traits. Variation in R may thus reflect both environmental (e.g., microhabitat experienced during development) and genetic factors (e.g., load of deleterious alleles; pathogen resistance; possession of locally beneficial alleles).

Let us write fitness as a function $W(x_1, x_2, \ldots, x_n)$ of the fitness components x_i. Suppose that an optimal strategy for a given individual involves a nonzero investment in the fitness components x_i and x_j (we will consider the case of zero investment in a fitness component below). It follows that the marginal fitness increase from investing in each component must be equal at this optimum—otherwise it would pay to reallocate resources toward the component with the highest marginal fitness increase (the "marginal advantage theorem"; Lloyd 1988). Formally, this means that for any fitness components x_i and x_j with $u_i, u_j > 0$ at the optimum, we must have

$$\left.\frac{\partial W}{\partial u_i}\right|_{u_i=u_i^*} = \left.\frac{\partial W}{\partial u_j}\right|_{u_j=u_j^*}. \qquad (1)$$

Note that these derivatives are evaluated at the optimum values of u_i and u_j, respectively, denoted here as u_i^* and u_j^*. What this implies about optimal resource allocation depends on the functional form of W, which reflects the type of trade-off at hand. If the fitness components can be measured in the same currency (e.g., the number of offspring produced, assuming offspring have equal reproductive value), fitness can be expressed as

$$W(x_1, x_2, \ldots, x_n) = x_1 + x_2 + \cdots + x_n. \qquad (2)$$

For example, consider the trade-off between producing intra-versus extrapair offspring. Here, the fitness value of a given offspring to its father may be independent of how many offspring of the other type he sires. Similar considerations may also apply to production of sons versus daughters, to production through male versus female function (in hermaphrodites), to production of offspring versus other relatives (in cooperative breeders), or various combinations of these possibilities.

Substituting equation (2) into equation (1), we learn that if fitness is additive and an optimal strategy involves nonzero investments in both x_i and x_j, then

$$\left.\frac{dx_i}{du_i}\right|_{u_i=u_i^*} = \left.\frac{dx_j}{du_j}\right|_{u_j=u_j^*}. \qquad (3)$$

In other words, at the optimal allocation, the marginal absolute increase in each fitness component with respect to investment must be equal.

Alternatively, a trade-off may exist between multiplicative fitness components, that is, those whose product is fitness, as

$$W(x_1, x_2, \ldots, x_n) = x_1 x_2 \cdots x_n. \qquad (4)$$

For example, male fitness might be expressed as: (expected life span) * (proportion of life spent reproductively active) *

(mating rate during this time) * (number of fertilizable eggs per mating) * (expected paternity share) * (recruitment probability per offspring) * (reproductive value per offspring recruited). Substituting equation (4) into equation (1) yields

$$\left.\frac{dx_i}{du_i}\right|_{u_i=u_i^*} \prod_{k\neq i} x_k\left(u_k^*\right) = \left.\frac{dx_j}{du_j}\right|_{u_j=u_j^*} \prod_{k\neq j} x_k\left(u_k^*\right). \tag{5}$$

Thus, optimal allocation equalizes the marginal increase of each fitness component x_i with respect to investment, weighted by the product of all other components. For a given pair of fitness components x_i and x_j, this can also be expressed as

$$\frac{1}{x_i\left(u_i^*\right)}\left(\left.\frac{dx_i}{du_i}\right|_{u_i=u_i^*}\right) = \frac{1}{x_j\left(u_j^*\right)}\left(\left.\frac{dx_j}{du_j}\right|_{u_j=u_j^*}\right). \tag{6}$$

In other words, the *marginal proportional increase* (known in economics as "elasticity") must be equal for each fitness component that receives a nonzero investment. This means that a small absolute increase of investment u_i in any such fitness component (while holding other investments unchanged) must lead to the same proportional increase in that component, and hence in fitness itself. For example, increasing mating rate by 1% will increase fitness by 1%, all else being equal. Each side of equation (6) can be interpreted as a selection gradient, that is, as a proportional derivative of the fitness component with respect to investment:

$$S_i(u_i) = \frac{x'_i(u_i)}{x_i(u_i)}. \tag{7}$$

Hence, equation (6) says that if an optimal strategy invests in both x_i and x_j, then $S_i(u_i) = S_j(u_j)$ at the optimum. Similarly, if an optimal strategy invests in x_i but not x_j (i.e., $u_i > 0$ but $u_j = 0$), then $S_i(u_i) \geq S_j(0)$.

Interestingly, previous theory on honest signaling has sometimes combined aspects of multiplicative and additive models. For instance, Iwasa et al. (1991) and Johnstone et al. (2009) assume that fitness is given by the product of exponential components. This product becomes additive on the logarithmic scale, allowing the authors to apply mathematical techniques that were developed for the case of additive trade-offs. In contrast, the arguments we develop below apply directly to multiplicative trade-offs without the need to assume particular functional forms. Instead, we assume that trade-offs arise due to reliance on a common limiting resource.
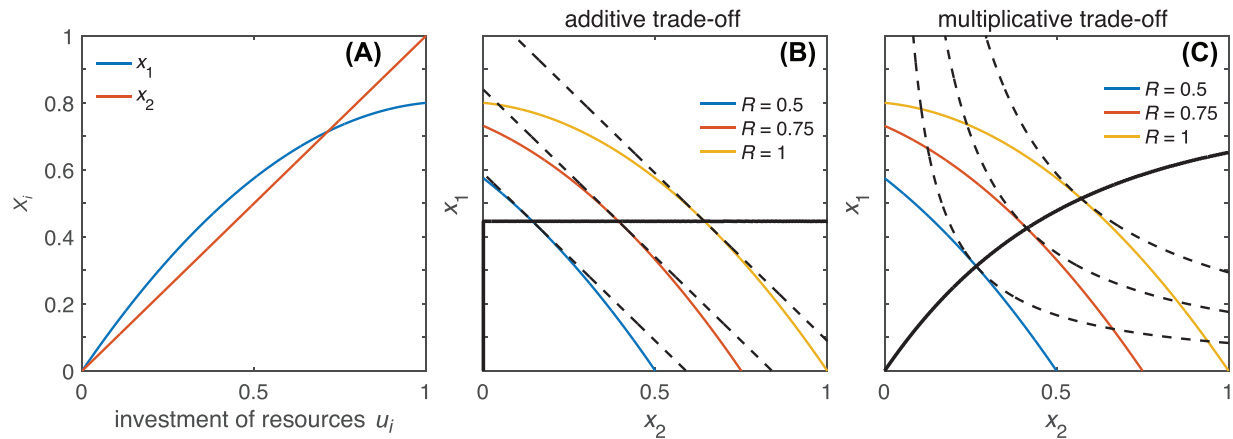
Signal Honesty

To link these general considerations to sexual signaling, let us suppose that fitness is a multiplicative function of mating success x_1 and viability x_2. This might hold approximately for males in species with no paternal investment in offspring, such that a male's mating success is a good proxy for his reproductive success. Furthermore, let us assume that investment u_1 is translated into mating success x_1 via expressing a signal trait t_1, such that t_1 increases monotonically with u_1 and x_1 increases monotonically with t_1. Thus, t_1 provides reliable information about R whenever u_1 increases monotonically with R, allowing us to use the latter relationship as our criterion for signal honesty. For simplicity, here we focus on the case where signal investment reduces viability indirectly, by consuming resources that would otherwise be invested in viability. However, as we show in the appendix, an analogous argument applies when signals also reduce viability directly (e.g., by attracting predators).

We begin with a heuristic argument to show why honest signaling should not surprise us (for formal proofs, see the appendix). From equation (7), we can see that if we hypothetically increased x_2 while its derivative x'_2 remained constant, then S_2 would decrease. This represents weaker selection for investing in u_2. Now, at the optimum, we know from equation (6) that selection favors investments in u_1 and u_2 equally. Consequently, if the derivative $x'_2(u_2)$ does not vary too strongly near the optimum, then there is an incentive to accompany any small increase in u_2 (due to an increase in R) with a small (not necessarily equal) increase in u_1. This selects for u_1 to strictly increase with R, ensuring honesty, at least if the signaling strategy is constrained to be a continuous function of resource level. Even if discontinuous strategies are permitted, the optimal investment u_1 is an increasing function of resources R whenever $S_2(u_2)$ is a strictly decreasing function of investment in viability (proof in appendix). On the other hand, if a small increase in x_2 is accompanied by a proportionally greater increase in x'_2, formally, if

$$\frac{x''_2(u_2)}{x'(u_2)} > \frac{x'_2(u_2)}{x_2(u_2)}, \tag{8}$$

then S_2 increases despite increasing x_2. This increase in S_2 creates an incentive to direct further investments to u_2 rather than to u_1. As a result, u_1 may decrease with R, undermining honesty. Because both x_2 and x'_2 are assumed to be positive (i.e., x_2 is a positive-valued trait that increases with investment), condition (8) cannot be satisfied if x_2 is linear ($x''_2 = 0$) or decelerating ($x''_2 < 0$). This guarantees honesty under a wide range of biologically plausible conditions, namely, those where viability does not increase with investment in an accelerating manner. As illustrated in Figure 1, the same does not hold if x_1 and x_2 are subject to an additive rather than a multiplicative trade-off. Condition (8) is necessary but not sufficient for the breakdown of honesty, because the effect of a moderately accelerating function $x_2(u_2)$ may be outweighed by S_2's tendency to decrease with increasing x_2 (all else being equal), or by an accelerating function $x_1(u_1)$. So, to undermine honesty, the acceleration of viability as a function of investment needs to be sufficiently strong (as, e.g., in Fig. 2A).

Figure 1. Example where a multiplicative (but not an additive) trade-off ensures signal honesty. (A) Fitness components x_1 (mating success) and x_2 (viability) as functions of the absolute amount of resources allocated to them, shown here as $x_1 = 1.5u_1 - 0.7u_1^2$ and $x_2 = u_2$. We assume that mating success is an increasing function of signal strength, which in turn increases with investment u_1. (B) Additive trade-off between x_1 and x_2, with fitness given by $W = x_1 + x_2$. Colored lines represent achievable trait combinations (phenotype sets) for individuals with fixed resource levels (see legend within panel). Dashed lines are fitness isoclines that connect hypothetical trait combinations yielding the same fitness. For each colored line (resource level), the optimal allocation occurs at the point where it touches the highest fitness isocline, corresponding to the highest achievable fitness for a given resource budget. The solid black line connects optimal trait combinations that correspond (from left to right) to increasing resource budgets. The solid black line's flat slope shows that individuals of any quality have the same mating success (hence signal size); that is, there is no honest signaling of quality. This result obtains because there is only a single point (at $u_1 = 0.36$, corresponding to $x_1 = 0.45$) at which both curves shown in panel A have matching slopes, as they must at equilibrium according to equation (3). (C) Multiplicative trade-off between x_1 and x_2, with fitness given by $W = x_1 x_2$. Lines have the same meaning as in panel B. The solid black line's consistently positive slope indicates that signaling is honest, that is, higher quality individuals have bigger signals and hence higher mating success.
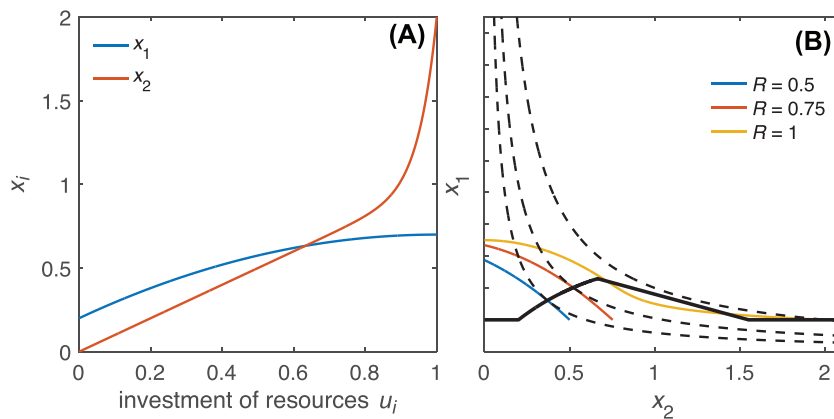
Figure 2. Example where a multiplicative trade-off (i.e. fitness given by $W = x_1 x_2$) fails to guarantee signal honesty. (A) Fitness components x_1 (mating success) and x_2 (viability) as functions of the absolute amount of resources allocated to them, shown here as $x_1 = 0.2 + u_1 - 0.5u_1^2$ and $x_2 = u_2 + u_2^{20}$. We assume that mating success is an increasing function of signal strength, which in turn increases with investment u_1. (B) Colored lines represent achievable trait combinations for individuals with fixed resource levels (see legend within panel). For each colored line (resource level), the optimal allocation occurs at the point where it touches the highest fitness isocline (dashed line). The solid black line connects optimal trait combinations that correspond (from left to right) to increasing resource budgets. The solid black line shows a negative slope over part of its range, indicating that higher quality individuals have smaller signals and lower mating success; that is, signaling is not honest. This outcome arises because the strongly accelerating returns of viability at high investment levels (upturned red curve on right-hand side of panel A) induce high-quality individuals to shift allocation toward viability at the expense of mating success.
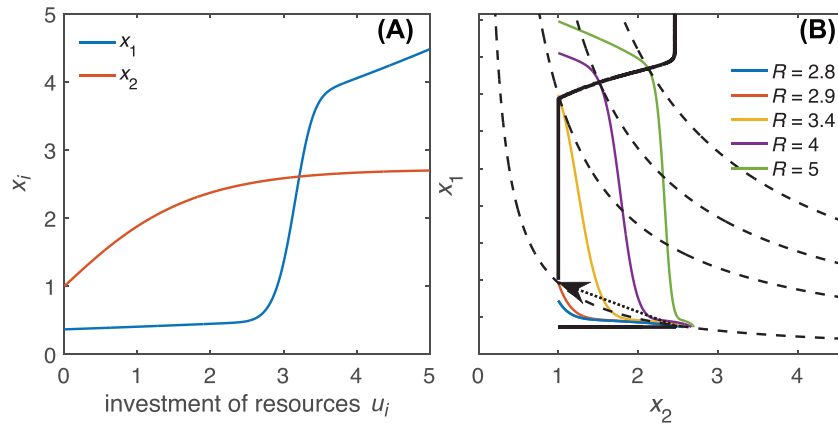
Figure 3. Example of honest signaling arising from a complicated allocation pattern under a multiplicative trade-off (i.e., fitness given by $W = x_1 x_2$). (A) Fitness components x_1 (mating success) and x_2 (viability) as functions of the absolute amount of resources allocated to them, shown here as $x_1 = \exp[0.1u_1 + \mathrm{erf}[3(u_1 - 3)]]$ (where $\mathrm{erf}[\cdot]$ is the Gauss error function) and $x_2 = \exp[1 - \exp[-u_2]]$. We assume that mating success is an increasing function of signal strength, which in turn increases with investment u_1. (B) Colored lines represent achievable trait combinations for individuals with fixed resource levels (see legend within panel). For each colored line (resource level), the optimal allocation occurs at the point where it touches the highest fitness isocline (dashed line). The solid black line connects optimal trait combinations that correspond (from bottom left to top right) to increasing resource budgets. At low resource levels, individuals do not invest in signaling (indicated by the flat slope at $x_1 = 0.37$, which corresponds to $u_1 = 0$). Optimal strategies then take an abrupt turn: after a brief stretch where x_1 increases while x_2 remains constant (up to $R \approx 2.9$, not discernible at this resolution) follows a discontinuity, where x_1 jumps upward and x_2 downward (marked with an arrow). This is followed by a further period of increasing x_1 with stationary x_2 (up to $R \approx 3.4$), after which x_1 and x_2 again increase together. Despite this complexity, larger signals always imply higher quality, meaning that signaling is honest (as will always be the case when S_2 is strictly decreasing: see the appendix).

To get an intuitive sense of why the trade-off's multiplicative nature matters, consider how the curvature of the fitness isoclines in Figure 1B, C affects their contact points with the "phenotype sets" (possible trait combinations) associated with different qualities. (This representation is inspired by fitness set theory [Levins 1962].) If phenotype sets are of similar shape (indicating that individuals of different quality face quantitatively similar trade-offs), then high-quality individuals reach their "target isocline" at higher values of x_1 than do low-quality individuals. That is, high-quality individuals enjoy higher mating success than low-quality individuals under their respective optimal strategies. By contrast, to undermine honesty, phenotype sets must differ in shape such that either (i) one of the lower phenotype sets extends upward to touch its target isocline at a relatively high value of x_1 (as in Fig. 1B) or (ii) one of the higher phenotype sets extends to the right to touch its target isocline at a relatively low value of x_1 (as in Fig. 2B). Such outcomes are harder to attain (i.e., require more dramatic deformations of phenotype sets) when fitness isoclines are bent in a downward-convex manner as in Figure 1C (in contrast to Fig. 1B), making their high-x_1 and low-x_1 regions harder to reach. Even so, deviations from honesty remain possible if phenotype sets differ sufficiently strongly in shape. For example, when high values of u_2 yield dramatic increases in x_2 (Fig. 2A), this manifests as phenotype sets of high-quality

individuals extending strongly to the right. In other words, if enhancing viability gets cheaper toward high investment levels, then high-quality individuals may shift their allocation to take advantage of this, even at the expense of reduced signal size (hence mating success).

We note above that if fitness is multiplicative and $S_2(u_2)$ is strictly decreasing, then signaling will always be honest, in the sense that larger signals indicate higher quality (proof in the appendix). Even if this condition is met, however, the relationship between quality and investment patterns can be quite nuanced. First, if individuals gain some mating success even without signaling, then individuals with fewer resources may be best off not signaling at all (Fig. 3). In such cases, larger signals still reliably indicate higher quality, but the absence of a signal merely signifies that an individual falls into the lower range of quality. Second, investment in viability u_2 may depend nonmonotonically on quality, even when signaling investment u_1 is strictly increasing (Fig. 3). This can occur, for instance, when there is a sharp increase in the slope of the relationship between investment and mating success. Such changes in slope can even select for individuals to increase signal investment more than proportionally to gains in resources. This results in a discontinuity (cf. Clifton et al. 2016), where signal investment jumps upward and viability investment jumps downward (Fig. 3).

Discussion

Here, we argue that when mating success and viability combine multiplicatively to determine fitness, then honest secondary sexual signals can evolve under broad conditions. As individual resource budgets increase, the adaptive need to keep multiplicative fitness components in balance is usually (but not always) best served by increasing investment in each. If this is the case—and assuming individuals use their optimal allocations—then an individual's resource budget (i.e., quality) can be inferred from its investment in any given fitness dimension, and in particular from signal traits that reflect investment in mating success.

The basic idea behind this reasoning can be illustrated with the following metaphor. Consider an old mechanical balance that is partly hidden from view, such that one can only see the weight carried by one of its arms. The partial view then gives full information about the total weight carried, provided we have reason to believe that the arms are, in fact, balanced. (To avoid the unwanted impression that "balanced" necessarily means "equal" here, we may allow that the balance has arms of different [known] lengths.)

In the context of sexual signaling, the reason for expecting balanced investments is rooted in the Darwinian postulate that natural selection is an optimizing process, which (in the long run) tends to shape organisms toward phenotypes that maximize their individual (inclusive) fitness (e.g., Fromhage and Jennions 2019). If fitness is multiplicative, maximizing an organism's fitness is analogous to maximizing the area of a rectangle, where the rectangle's length and width correspond to fitness dimensions in which an individual may invest. If the resource costs of increasing length or width by the same absolute amount are equal, then the most efficient strategy is to increase investment in whichever dimension is currently smallest. Even if length and width differ in cost, optimal strategies must roughly balance investments in each dimension.

Its adaptationist foundation sets our model apart from the classic work by Van Noordvijk and de Jong (1986), which showed that positive covariance between life-history traits readily arises when individual resource budgets exhibit high variance relative to variance in allocation. Their finding has been called the "big house, big car effect" (Reznick et al. 2000), based on an analogy in economics: "if the budget is fixed, people spending more on housing should spend less on cars. In fact, the amount of expendable income is variable, and in many situations positive correlations are observed between the per-family expenses on housing and cars" (van Noordwijk and de Jong 1986, p. 141). Yet this analogy, such as Van Noordvijk and de Jong's model, leaves open why resources should be allocated in this way. Their model simply assumes that resource acquisition is independent of relative allocation, thereby excluding the possibility that individuals adjust their allocation to maximize fitness given their budget. By contrast, our present model provides an adaptive rationale for the "big house, big car effect" as it pertains to sexual signaling and viability.

Although Grafen (1990) interpreted individual quality as an abstract property with no explicit mechanistic link to trait expression, here we have interpreted quality as a measure of an individual's resource budget. This formulation accommodates the plausible view that, say, producing an additional eye spot costs the same absolute amount of resources for peacocks of any quality, while amounting to a higher proportional investment for low-quality peacocks. From this natural link between absolute and proportional investments arises a dependence between individual quality and signaling costs, which—without ad hoc assumptions about its shape—justifies the intuitive notion that greater resource budgets will usually lead to greater optimal spending on sexual signals.

Let us now turn to the exceptions from the prediction that higher quality individuals should produce stronger signals. We may ask: how plausible is it that viability is an accelerating function of investment, to the extent that high-quality individuals are incentivized to neglect investment in mating success? The answer potentially depends on which aspect of viability we consider. For viability traits that are conceived of as probabilities (e.g., survival until the onset of reproduction), a diminishing rather than accelerating return on investment seems likely, at least near the theoretical maximum of 1. Biologically, we may expect that, if some causes of mortality are harder to avoid then others, then avoiding the easy ones should take priority. Thus, as an individual increases its investment in survival, each further increment may be harder to attain (i.e., require more investment) than the previous one. On the other hand, for the reproductive life span in iteroparous species—another important aspect of viability—the lack of a theoretical upper limit gives less reason to argue against accelerating returns a priori. It would therefore be desirable to study such relationships empirically, ideally by measuring or manipulating investment in survival and recording associated changes in life span.

For simplicity, we have focused on fitness trade-offs from the perspective of the signaler, while ignoring the evolution of receiver preferences. Preferences for larger signals could conceivably be favored if quality (i.e., resource level) is partly heritable, because both male and female offspring would benefit from higher resource availability (i.e., a "good genes" model: Iwasa et al. 1991; Kuijper et al. 2012; Dhole et al. 2018). Moreover, the broad conditions for honest signaling in our model suggest that honesty should be robust to substantial variation in receiver behavior, as long as there remains a positive correlation between signal size and mating success. We consequently predict that the

insights of our model will hold up in coevolutionary models of signal and preference.

A further limitation of our model is that, for simplicity, it does not allow for individuals to adjust allocation repeatedly during their lifetime, for example, in response to parasitism or other adverse life events. Yet such flexible adjustment appears to operate in several empirically documented cases of "dishonest" sexual signaling. For example, parasitized *Drosophila* males increase their courtship activity (Polak and Starmer 1998) and food-deprived stickleback males exhibit brighter nuptial coloration (Candolin 1999), presumably in a "terminal effort" to maximize reproduction in what little time they have left. Such males then no longer face the mating-viability trade-off as modeled here, as their fitness gains in the short term (which is all that counts for them) may require little or no investment in viability.

Signal traits that are deemed honest and costly are often loosely labeled as handicaps without worrying about the term's history and connotations. Yet appeals have been made to usher the handicap hypothesis and its associated terminology into an "honorable retirement" (Getty 2006; Penn and Számadó 2020). There are two main arguments for this. First, the handicap hypothesis's central prediction (that costs are necessary to explain the evolution of signal reliability) has failed the test of time (Hurd 1995; Getty 1998; Számadó 1999; Lachmann et al. 2001; Higham 2013; Penn and Számadó 2020). In particular, honest signaling does not strictly require realized costs at equilibrium (as in our model); rather, it suffices that dishonesty is costly. Second, the handicap metaphor may owe much of its popularity to Zahavi's paradoxical sounding and indeed misleading rhetoric about selection for waste rather than efficiency (Zahavi 1981, 1987; Zahavi and Zahavi 1997). According to Penn and Számadó (2020, p 12), in the 1980s "Finding a model that supports the Handicap Principle became a theoretical challenge in evolutionary biology comparable to constructing a perpetual motion machine…"; that is, highly coveted yet deemed impossible. This climate generated much attention when Grafen (1990) claimed to have vindicated Zahavi (1975), thus making Zahavi's paper (and subsequent book) citation classics of the biological and social sciences in spite of their questionable aspects. From a modern standpoint, the "waste" metaphor appears to offer little of value. At the level of individual organisms, an optimal strategy might trade off one fitness component against another, but this could hardly be called "waste" if the result it to maximize fitness. On the other hand, competitive traits may very well reduce population fitness; however, this is a very general phenomenon not limited to sexual signaling (e.g., the height of trees can also be viewed in this light). Whether we can move beyond this potential source of confusion may depend, in part, on biologists' willingness to abandon a memorable metaphor. To aid this process, here we have advo-cated an alternative metaphor: of costly signals as balanced investments rather than handicaps.

Appendix

Proof that signaling is honest

Here, we characterize the relationship between an individual's signal size (a strictly increasing function of its investment u_1) and its resource level R. We assume throughout that the selection gradient $S_2(u_2)$ on investment in viability is a strictly decreasing function of the current investment level u_2. Note that

$$S'_2(u_2) = \frac{x''_2(u_2)}{x_2(u_2)} - S_2(u_2)^2. \tag{A1}$$

The function $S_2(u_2)$ is strictly decreasing if $S'_2(u_2) < 0$. A sufficient (but not necessary) condition is that $x_2(u_2)$ is concave (i.e. $x''_2(u_2) \leq 0$).

We will prove the following theorem:

(**Honest signaling theorem**): If $S_2(u_2)$ is strictly decreasing, then there is a threshold resource level R^* such that the optimal investment in signaling is

(i) *zero for all $R < R^*$, and*
(ii) *a strictly increasing function of R for all $R > R^*$.*

The cases $R^* = 0$ and $R^* = \infty$ are not excluded. If $R^* = 0$, then optimal signaling investment always increases reliably with the resource level (i.e., an individual's signal size gives complete information about their resource level; this is sometimes known as a "separating equilibrium"). If $R^* = \infty$, then individuals do not signal. For $0 < R^* < \infty$, individuals with fewer resources do not signal, whereas those with greater resources increase their signal size along with their resource level. We will first prove the above theorem and then consider some factors influencing R^*.

We note that for a given resource level R, there may be more than one optimal strategy $u_1^*(R)$. The statement that optimal investment is strictly increasing should consequently be interpreted as follows: If $R_1 < R_2$, then for all optimal strategies $u_1^*(R_1)$ and $u_1^*(R_2)$, we have $u_1^*(R_1) < u_1^*(R_2)$.

Optimal investment is a nondecreasing function of resource level

We first show that optimal investment in signaling is a nondecreasing function of the resource level R. It is helpful to first prove a lemma about the gains in viability with increasing investment. Suppose $\Delta > 0$ is a fixed constant and let $f_\Delta(u_B)$ be the proportional gain in viability when increasing investment u_2 from a baseline of u_B to a new level of $u_B + \Delta$:

$$f_\Delta(u_B) = \frac{x_2(u_B + \Delta)}{x_2(u_B)}. \tag{A2}$$

Note that

$$f'_\Delta(u_B) = \frac{x'_2(u_B + \Delta)}{x_2(u_B)} - \frac{x_2(u_B + \Delta)\,x'_2(u_B)}{x_2(u_B)^2}$$

$$= \frac{x_2(u_B + \Delta)}{x_2(u_B)}(S_2(u_B + \Delta) - S_2(u_B)). \tag{A3}$$

Because S_2 is strictly decreasing and $\Delta > 0$, we have $f'_\Delta(u_B) < 0$. Hence, f_Δ is also a strictly decreasing function. This means that the proportional increase in viability when increasing investment by a fixed absolute amount Δ is a strictly decreasing function of baseline investment u_B.

Now, suppose there exist $R_1 < R_2$ and optimal strategies $u_1^*(R_1)$ and $u_1^*(R_2)$ such that $u_1^*(R_1) > u_1^*(R_2)$. We will show that an individual with resource level R_2 would achieve higher fitness by investing $u_1^*(R_1)$ in mating success than by investing $u_1^*(R_2)$, which contradicts the optimality of $u_1^*(R_2)$. Let $\Delta = u_1^*(R_1) - u_1^*(R_2) > 0$. Because $R_1 < R_2$ and $f_\Delta(u_B)$ is strictly decreasing, we have

$$\frac{x_2(R_2 - u_1^*(R_2))}{x_2(R_2 - u_1^*(R_1))} = f_\Delta(R_2 - u_1^*(R_1)) < f_\Delta(R_1 - u_1^*(R_1))$$

$$= \frac{x_2(R_1 - u_1^*(R_2))}{x_2(R_1 - u_1^*(R_1))}. \tag{A4}$$

Further, if $u_1^*(R_1)$ is an optimal strategy for an individual with resource level R_1, then it must yield fitness at least as high as any alternative strategy $\hat{u}_1$. In particular, when $\hat{u}_1 = u_1^*(R_2)$, we must have

$$x_1(u_1^*(R_1))\,x_2(R_1 - u_1^*(R_1)) \ge x_1(u_1^*(R_2))\,x_2(R_1 - u_1^*(R_2)). \tag{A5}$$

Rearranging this equation yields

$$\frac{x_2(R_1 - u_1^*(R_2))}{x_2(R_1 - u_1^*(R_1))} \le \frac{x_1(u_1^*(R_1))}{x_1(u_1^*(R_2))}. \tag{A6}$$

Lastly, combining the inequalities (A4) and (A6) and rearranging, we have

$$x_1(u_1^*(R_2))\,x_2(R_2 - u_1^*(R_2)) < x_1(u_1^*(R_1))\,x_2(R_2 - u_1^*(R_1)). \tag{A7}$$

In other words, the fitness of an individual with resource level R_2 is higher when playing $u_1^*(R_1)$ than when playing $u_1^*(R_2)$. However, this contradicts the optimality of $u_1^*(R_2)$. Therefore, optimal investment in signaling is a nondecreasing function of an individual's resource level.

Nonzero optimal investments at different resource levels can never be equal

Next, let us suppose that for some $R_1 < R_2$ there exist optimal strategies $u_1^*(R_1)$ and $u_1^*(R_2)$ such that $u_1^*(R_1) = u_1^*(R_2) > 0$.

The corresponding optimal investments in viability are $R_1 - u_1^*(R_1)$ and $R_2 - u_1^*(R_2)$, respectively. From equations (6) and (7) in the main text, we have

$$S_2(R_1 - u_1^*(R_1)) = S_1(u_1^*(R_1)) = S_1(u_1^*(R_2))$$

$$= S_2(R_2 - u_1^*(R_2)). \tag{A8}$$

However, because $R_1 - u_1^*(R_1) < R_2 - u_1^*(R_2)$, this contradicts the assumption that S_2 is strictly decreasing. Therefore, if $R_1 \ne R_2$, then nonzero optimal investments $u_1^*(R_1)$ and $u_1^*(R_2)$ can never be equal. Combined with the previous result, this means that if there exists a resource level $\hat{R}$ at which signaling investment is positive, then $u_1^*(R)$ is strictly increasing for all $R \ge \hat{R}$.

We can now combine the above results to prove the honest signaling theorem. Let $R^* \in [0, \infty]$ be the smallest resource level such that $R \le R^*$ whenever $u_1(R) = 0$ is optimal (i.e., $R^* = \sup\{R : u_1(R) = 0 \text{ is optimal}\}$). Because $u_1^*(R)$ is nonnegative and nondecreasing, we must have $u_1^*(R) = 0$ for all $R < R^*$. Further, because $u_1^*(R)$ is positive for all $R > R^*$, it is a strictly increasing function on the same interval.

What determines R^*?

Here, we note some facts about the threshold resource level R^* above which individuals signal:

(i) Individuals always signal (i.e., $R^* = 0$) if and only if $S_1(0) \ge S_2(0)$.

(ii) If $S_1(u_1) < S_2(u_2)$ for all u_1 and u_2, then individuals never signal (i.e., $R^* = \infty$).

(iii) If there exists a resource value $\hat{R}$ such that $S_1(0) > S_2(\hat{R})$, then individuals signal for all $R \ge \hat{R}$ (i.e., $R^* \le \hat{R}$). In particular, if the proportional marginal gains in viability tend to zero as investment increases (i.e., $S_2(u_2) \to 0$ as $u_2 \to \infty$), then there is a resource level above which individuals signal (i.e., $R^* < \infty$).

Propositions (ii) and (iii) are fairly straightforward. For (i), note that because S_2 is strictly decreasing, $S_1(0) \ge S_2(0)$ implies that $S_1(0) > S_2(R)$ for all $R > 0$. Consequently, there can be no optimum of the form $u_1^*(R) = 0$ and $u_2^*(R) = R$. Conversely, if $S_1(0) < S_2(0)$ then, by continuity, there is a value $R_\delta > 0$ such that $S_1(u_1) < S_2(u_2)$ for all $u_1 + u_2 < R_\delta$. This implies that $u_1^*(R) = 0$ is optimal for all $R \le R_\delta$.

When viability depends directly on signal size

In the main text, we assume that signal investment reduces viability indirectly, by consuming resources that would otherwise be invested in viability. However, signals might also reduce viability directly (e.g., by attracting predators). In this case, viability x_2 is a function of both signal investment u_1 and viability investment u_2.

Shifting resources from u_2 to u_1 reduces viability via two distinct mechanisms: reduced viability investment and increased signal size. We can then write fitness as

$$W = x_1(u_1)\,x_2(u_1, u_2). \qquad (A9)$$

We write $S_2(u_1, u_2)$ for the selection gradient that captures both of the viability effects associated with infinitesimal changes in u_2:

$$S_2(u_1, u_2) = \frac{1}{x_2}\left(\frac{\partial x_2}{\partial u_2} - \frac{\partial x_2}{\partial u_1}\right). \qquad (A10)$$

Analogously to above, signaling is guaranteed to be honest if, for fixed u_1, $S_2(u_1, u_2)$ is a strictly decreasing function of u_2:

$$\frac{\partial}{\partial u_2}S_2(u_1, u_2) = \frac{1}{x_2}\left(\frac{\partial^2 x_2}{\partial u_2^2} - \frac{\partial^2 x_2}{\partial u_1 \partial u_2} - \frac{\partial x_2}{\partial u_2}\,S_2(u_1, u_2)\right) < 0. \qquad (A11)$$

In other words, for a fixed signal size, the proportional increase in viability when moving resources from u_1 to u_2 is a decreasing function of u_2.

The proof is analogous to above. First, let us define

$$f_\Delta(u_1, u_2) = \frac{x_2(u_1 - \Delta, u_2 + \Delta)}{x_2(u_1, u_2)}. \qquad (A12)$$

We then have

$$\frac{\partial}{\partial u_2}f_\Delta(u_1, u_2) = \frac{x_2(u_1 - \Delta,\ u_2 + \Delta)}{x_2(u_1, u_2)}\left(g_{u_1, u_2}(\Delta) - g_{u_1, u_2}(0)\right), \qquad (A13)$$

where

$$g_{u_1, u_2}(\Delta) = \frac{x_2^{(0,1)}(u_1 - \Delta, u_2 + \Delta)}{x_2(u_1 - \Delta, u_2 + \Delta)}. \qquad (A14)$$

Here, we write $f^{(0,1)}$ to represent the partial derivative of a binary function f with respect to its second argument, while holding the first argument fixed. Now note that

$$\frac{\partial}{\partial \Delta}g_{u_1, u_2}(\Delta) = S_2^{(1,0)}(u_1 - \Delta, u_2 + \Delta), \qquad (A15)$$

which is negative by assumption. Hence, $g_{u_1, u_2}(\Delta) - g_{u_1, u_2}(0) < 0$ and so $\frac{\partial}{\partial u_2}f_\Delta(u_1, u_2) < 0$. In other words, the proportional increase in viability when reallocating Δ resources from u_1 to u_2 is a decreasing function of u_2. The rest of the proof closely follows the one above.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS
L.F. had the idea and wrote the first draft. J.M.H. wrote the Appendix and contributed through discussion of ideas and writing.

LITERATURE CITED
Andersson, M. 1994. Sexual selection. Princeton Univ. Press, Princeton, NJ.

Candolin, U. 1999. The relationship between signal quality and physical condition: is sexual signalling honest in the three-spined stickleback? Anim. Behav. 58:1261–1267.

Clifton, S. M., R. I. Braun, and D. M. Abrams. 2016. Handicap principle implies emergence of dimorphic ornaments. Proc. R. Soc. B Biol. Sci. 283:1–6.

Dhole, S., C. A. Stern, and M. R. Servedio. 2018. Direct detection of male quality can facilitate the evolution of female choosiness and indicators of good genes: evolution across a continuum of indicator mechanisms. Evolution 72:770–784.

Dougherty, L. R. 2021. Meta-analysis reveals that animal sexual signalling behaviour is honest and resource based. Nat. Ecol. Evol. 5:688–699.

Fromhage, L., and M. D. Jennions. 2019. The strategic reference gene: an organismal theory of inclusive fitness. Proc. R. Soc. B Biol. Sci. 286:1–9.

Getty, T. 1998. Reliable signalling need not be a handicap - commentary. Anim. Behav. 56:253–255.

———. 2006. Sexually selected signals are not similar to sports handicaps. Trends Ecol. Evol. 21:83–8.

Grafen, A. 1990. Biological signals as handicaps. J. Theor. Biol. 144:517–546.

Higham, J. P. 2013. How does honest costly signaling work? Behav. Ecol. 25:8–11.

Houston, A. I., J. M. McNamara, and M. L. Hernandez. 2003. How does an individual's optimal behaviour depend on its quality? An analysis based on relative ability. Evol. Ecol. Res. 5:195–212.

Hurd, P. L. 1995. Communication in discrete action-response games. J. Theor. Biol. 174:217–222.

Iwasa, Y., A. Pomiankowski, and S. Nee. 1991. The evolution of costly mate preferences. II. The "handicap" principle. Evolution 45:1431–1442.

Johnstone, R. A., S. A. Rands, and M. R. Evans. 2009. Sexual selection and condition-dependence. J. Evol. Biol. 22:2387–2394.

Kirkpatrick, M. 1986. The handicap mechanism of sexual selection does not work. Am. Nat. 127:222–240.

Kuijper, B., I. Pen, and F. J. Weissing. 2012. A guide to sexual selection theory. Annu. Rev. Ecol. Evol. Syst. 43:287–311.

Lachmann, M., S. Szamado, and C. T. Bergstrom. 2001. Cost and conflict in animal signals and human language. Proc. Natl. Acad. Sci. USA 98:13189–94.

Levins, R. 1962. Theory of fitness in a heterogeneous environment. I. The fitness set and adaptive function. Am. Nat. 96:361–378.

Lloyd, D. G. 1988. A general principle for the allocation of limited resources. Evol. Ecol. 2:175–187.

Penn, D. J., and S. Számadó. 2020. The Handicap Principle: how an erroneous hypothesis became a scientific principle. Biol. Rev. 95:267–290.

Polak, M., and W. T. Starmer. 1998. Parasite-induced risk of mortality elevates reproductive effort in male *Drosophila*. Proc. R. Soc. B Biol. Sci. 265:2197–2201.

Reznick, D., L. Nunney, and A. Tessier. 2000. Big houses, big cars, superfleas and the costs of reproduction. Trends Ecol. Evol. 15:421–425.

Somjee, U. 2021. Positive allometry of sexually selected traits: do metabolic maintenance costs play an important role? BioEssays 43:1–13.

Számadó, S. 1999. The validity of the handicap principle in discrete action-response games. J. Theor. Biol. 198:593–602.

van Noordwijk, A. J., and G. de Jong. 1986. Acquisition and allocation of resources - their influence on variation in life history tactics. Am. Nat. 128:137–142.

Zahavi, A. 1975. Mate selection-a selection for a handicap. J. Theor. Biol. 53:205–14.

———. 1981. Natural selection, sexual selection and the selection of signals. Pp. 133–138 *in* G. G. E. Scudder and J. L. Reveal, eds. Evolution today, Proceedings of the Second International Congress of Systematic and Evolutionary Biology. Carnegie-Mellon Univ. Press, Pittsburgh, PA.

———. 1987. The theory of signal selection and some of its implications. Pp. 305–327 *in* V. P. Delfino, ed. International Symposium on Biological Evolution. Adriatica Editrica Bari, Italy, France.

Zahavi, A., and A. Zahavi. 1997. The handicap principle: a missing piece of Darwin's puzzle. Oxford Univ. Press, New York.

Associate Editor: T. Connallon
Handling Editor: A. McAdam