

The Handicap Principle in Sexual Selection

Graham Bell

Evolution, Vol. 32, No. 4. (Dec., 1978), pp. 872-885.

Stable URL:

http://links.jstor.org/sici?sici=0014-3820%28197812%2932%3A4%3C872%3ATHPISS%3E2.0.CO%3B2-H

Evolution is currently published by Society for the Study of Evolution.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <u>http://www.jstor.org/journals/ssevol.html</u>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.

THE HANDICAP PRINCIPLE IN SEXUAL SELECTION

GRAHAM BELL

Department of Biology, McGill University, 1205 McGregor Avenue, Montreal, Quebec, Canada H3A 1B1

Received September 13, 1977. Revised January 10, 1978

The theory of the selection of epigamic color, structure and behavior was originally due to Darwin (1874) and was later elaborated by Fisher (1930) and by O'Donald (1962, 1963, 1967). It is well known and is discussed below. An alternative account of this type of sexual selection has been suggested recently by Zahavi (1975). The kernel of his idea is as follows. Suppose that in a certain species of animal the females are cryptically colored, while the males may be either cryptic or conspicuous; the argument will apply as well to structure or behavior as to coloration. Conspicuous males are assumed to survive less well than cryptic males; if this were not so, natural selection would suffice to explain the evolution of conspicuousness. Conspicuous males are thus the less fit, at least in terms of survival. But those conspicuous males which survive to reproduce despite this additional mortality must be, in other respects, extraordinarily fit. This fitness "in other respects" will be inherited without prejudice by their daughters and may be sufficient to counterbalance the lessened fitness of their male offspring. If this happens, then females will tend to evolve a preference to mate with conspicuous males, since those females which produce the most fit offspring will be favorably selected. The attraction of Zahavi's idea is its internal consistency and in particular the neatness with which the principal difficulty of sexual selection is revealed to be its motive force.

However, it is by no means self-evident that selection can act in this way, and the proposed mechanism has been severely criticised by Maynard Smith (1976), who was unable to find any combination of fitnesses and gene frequencies for which the handicap principle would produce effective sexual selection in the absence of the Fisher effect. Because Zahavi's idea would be of considerable interest if it could be shown to be feasible, the purpose of the present paper is a more extensive account of the behavior of models which incorporate the handicap principle. First, conditions that the handicap principle should be effective during the first generation of selection are obtained analytically. Secondly, the results of this analysis are compared with the results of computer simulation.

ANALYSIS

The analysis refers to autosomal loci in an infinite haploid population with discrete generations. At the first ("A") locus, male bearers of the A gene are conspicuous, whilst males bearing the a gene are cryptic; all females are cryptic, whatever their genotype. At the second ("B") locus, individuals bearing the B gene are the more fit, in some unspecified manner not connected with conspicuousness, whilst those bearing the b gene are the less fit. The frequency of the A gene is p, whilst the frequency of the B gene is q.

It is assumed that the requisite variation in female mating preference is generated at a third ("C") locus, where 'choosy' females bearing the c gene, prefer to mate with conspicuous males, whilst 'undiscriminating' females, bearing the c gene, display no preference between cryptic and conspicuous males. The frequency of the Cgene is r. However, the nature of this variation is not made explicit in the analysis where it is simply assumed that some genetic variance for mating preference exists. The axiom of the analysis is that if females of given genotype produce fitter offspring by mating with males of given genotype, then this type of mating will be favored by selection; and that if the products of this mating are fitter than the products of any

other type of mating, the two genotypes involved will increase in frequency.

Because the expression 'the A gene' is sex-limited, the table of fitnesses will be as follows:

| | Males: | | Females: |
|--------------|-----------|-----------|----------|
| | A (con- | a | (all |
| | spicuous) | (cryptic) | cryptic) |
| B (more fit) | w_{AB} | w_{aB} | w_{aB} |
| b (less fit) | w_{Ab} | w_{ab} | w_{ab} |

The frequency of the B gene after selection in cryptic and in conspicuous males respectively is then:

frequency of B gene in a males:

$$q_a = q w_{aB}/(qw_{aB} + (1 - q)w_{ab})$$

frequency of B gene in A males:
 $q_A = qw_{AB}/(qw_{AB} + (1 - q)w_{Ab})$

Thus, $q_A = q_a$ if $w_{AB}w_{ab} = w_{Ab}w_{aB}$; that is, if the genes at the *A* and *B* loci have multiplicative effects on fitness. For example, the genotype fitnesses might be as follows:

| | Males: | | Females: |
|--------------|-------------------------|------------------|------------------|
| | A (con- | a | (all |
| | spicuous) | (cryptic) | cryptic) |
| B (more fit) | lphaeta | α | α |
| b (less fit) | $\alpha \beta \epsilon$ | $\alpha\epsilon$ | $\alpha\epsilon$ |

Here, α is a 'baseline' fitness; β is the sexspecific effect on fitness at the *A* locus; ϵ is the locus-specific effect on fitness at the *B* locus. Since fitnesses are intended as rates of survival to reproductive age, we have $0 \leq \alpha, \beta, \epsilon \leq 1$. The effects on fitness at the two loci are multiplicative, so that $w_{AB}w_{ab} = w_{Ab}w_{aB} = \alpha^2\beta\epsilon$; hence $q_A =$ $q_a = \alpha q/(\alpha q + (1 - q)\alpha\epsilon)$. Because the frequency of the *B* gene after selection is the same in cryptic and in conspicuous males, selection can make no distinction between genotypes at the *A* locus by acting through genes at the *B* locus. Thus, the fate of conspicuous males will depend exclusively on selection acting at the A locus, as the result of which they will be eliminated from the population because of their lower fitness.

However, if gene effects are not multiplicative, this conclusion may not hold. If α , β and ϵ are additive gene effects, the table of fitnesses is as follows:

| Ma | les: | Females: |
|-----------|-----------|----------|
| A (con- | a | (all |
| spicuous) | (cryptic) | cryptic) |

B(more fit) $\alpha + \epsilon$ $\alpha + \beta + \epsilon \alpha + \beta + \epsilon$ b (less fit) α $\alpha + \beta$ $\alpha + \beta$ where again $0 \le \alpha, \beta, \epsilon \le 1$, and in addition $0 \le \alpha + \beta + \epsilon \le 1$. Since the effects on fitness are additive, $w_{AB}w_{ab} \ne w_{Ab}w_{aB}$ unless $\beta = 0$ or $\epsilon = 0$ or both. Conspicuous and cryptic males therefore differ with respect to the frequency of the B gene after selection, provided that neither allele at the B locus is fixed:

frequency of *B* gene in *a* males: $q_a = q(\alpha + \beta + \epsilon)/(\alpha + \beta + q\epsilon)$

frequency of *B* gene in *A* males:

$$\begin{cases}
(1) \\
q_A = q(\alpha + \epsilon)/(\alpha + q\epsilon)
\end{cases}$$

It is easily verified that $q_A > q_a$ if $\beta > 0$ and $\epsilon > 0$ and vice versa. Maynard Smith (1976) uses a scheme in which the fitnesses of genotypes amongst the males are: $w_{AB} = (1 - t); \ w_{ab} = (1 - s); \ w_{aB} = 1;$ $w_{Ab} = (1 - s)(1 - ut)$. This scheme is strictly multiplicative for u = 1, and corresponds to the additive scheme described above for u = 1/(1 - s); between these two points the fitnesses described are intermediate in character between the multiplicative and additive schemes.

| Zygote | Frequency | male fitness | female fitness |
|--------|----------------------|-----------------------------|-----------------------------|
| AB | $(1 + q_a)p/4$ | $\alpha + \epsilon$ | $\alpha + \beta + \epsilon$ |
| Ab | $(1 - q_a)p/4$ | α | $\alpha + \beta$ |
| aB | $(1 + q_a)(2 - p)/4$ | $\alpha + \beta + \epsilon$ | $\alpha + \beta + \epsilon$ |
| ab | $(1 - q_a)(2 - p)/4$ | $\alpha + \beta$ | $\alpha + \beta$ |

The mean fitness of the male offspring produced by this mating is therefore:

$$\bar{w}_{a,m} = \{ (\alpha + \epsilon)(1 + q_a)p + \alpha(1 - q_a)p + (\alpha + \beta + \epsilon)(1 + q_a)(2 - p) + (\alpha + \beta)(1 - q_a)(2 - p) \} \\ \times \frac{1}{4} \\ = \{ 2\alpha + (2 - p)\beta + (1 + q_a)\epsilon \}/2$$

Similarly, the mean fitness of the female offspring is:

$$\bar{w}_{o,f} = \{2\alpha + 2\beta + (1 + q_a)\epsilon\}/2$$

Assuming an equal primary sex ratio, the mean fitness of all offspring produced by this mating will be:

$$\bar{w}_{o,1} = \frac{1}{2} \bar{w}_{o,m} + \frac{1}{2} \bar{w}_{o,f} = \{ 4\alpha + (4 - p)\beta + 2(1 + q_a)\epsilon \} / 4$$

Performing the same calculation for the other three possible matings, we can write down the mean fitness of the offspring produced by a given mating:

$$(B \ \heartsuit \times a \ \eth): \bar{w}_{o,1} = \{4\alpha + (4 - p)\beta + 2(1 + q_a)\epsilon\}/4$$

$$(B \ \heartsuit \times A \ \eth): \bar{w}_{o,2} = \{4\alpha + (3 - p)\beta + 2(1 + q_A)\epsilon\}/4$$

$$(b \ \heartsuit \times a \ \eth): \bar{w}_{o,3} = \{4\alpha + (4 - p)\beta + 2q_a\epsilon\}/4$$

$$(b \ \heartsuit \times A \ \eth): \bar{w}_{o,4} = \{4\alpha + (3 - p)\beta + 2q_a\epsilon\}/4$$

$$(2)$$

A sufficient condition that *B* females should prefer to mate with *A* males is $\bar{w}_{o,2} > \bar{w}_{o,1}$; a sufficient condition that *b* females should prefer to mate with *A* males is $\bar{w}_{o,4} > \bar{w}_{o,3}$. A necessary (but not sufficient) condition for sexual selection is that at least one of these inequalities should be satisfied; a sufficient (but not necessary) condition is that both should be satisfied. Thus, sexual selection must follow if:

both:
$$\bar{w}_{o,2} > \bar{w}_{o,1}$$

and: $\bar{w}_{o,4} > \bar{w}_{o,3}$

This pair of inequalities is easily shown to reduce to the single sufficient condition:

$$2\epsilon(q_A - q_a) > \beta \tag{3}$$

This condition can be readily understood. The matings of A males produce the fittest offspring if the positive effect on fitness at the B locus (measured by ϵ), weighted by the excess of B genes amongst A males (measured by $q_A - q_a$), should exceed the negative effect at the A locus (measured by β). A factor of 2 appears because the expression of the A gene is limited to one of the two sexes. Because $(q_A - q_a)$ will usually be a small number, the handicap principle will be effective only if $\epsilon \gg \beta$; that is, if selection is nearly truncate, the effect on fitness of the A gene being much smaller than that of the b gene.

It is impracticable to carry the analytical argument beyond this point, since under the non-random mating implied by the preferece of C females for A males the genes at the three loci will evolve into linkage disequilibrium, so that the assumption that a fraction q_a of either choosy or of undiscriminating females bear the B gene will be invalidated. With linkage disequilibrium, the conditions for the spread of the A gene are quite different. If only the coupling genotypes BC and bcare present, then the offspring of A males will be the more fit if $\bar{w}_{o,2} > \bar{w}_{o,3}$, assuming that C females mate only with Amales, and this is equivalent to $2\epsilon(1 + \epsilon)$ $q_A - q_a$ > β . Since $q_A > q_a$, this is always satisfied if $2\epsilon > \beta$, a very lax condition. However, if only the repulsion genotypes bC and Bc are present, then the offspring of A males will be the more fit only if $\bar{w}_{o,4} > \bar{w}_{o,1}$, and a glance at equations (2) will show that this condition cannot be satisfied. Thus, the eventual fate of the A gene depends critically on the evolution of linkage disequilibrium within the population; the appearance of an excess of coupling genotypes will relax condition (3) to some extent.

SIMULATION

The analysis above shows conditions that the handicap principle should be effective during the first generation of selection, given an initial population in linkage equilibrium. The generation of linkage disequilibrium by the female mating preference makes it necessary to use numerical simulation on a computer in order to follow the course of selection in subsequent generations.

Programme design.—There are two sets of input variables. The first comprises the initial frequencies of the A, B and Cgenes, which I shall call the initial frequency set, $IfS = \{p_o, q_o, r_o\}$; unless otherwise stated, this was $IfS = \{0.001,$ 0.001, 0.001. The second comprises the three elements of the additive fitness scheme described above, which I shall call the fitness set, $FS = \{\alpha, \beta, \epsilon\}$. The remaining determinant of gene frequency is the set of mating rules, which by describing the preference of C females for Amales governs the weight which is to be given to each of the four possible matings with respect to the C locus in females and the A locus in males. Two different sets of rules were used in the simulations, the first describing a monogamous and the second a polygamous mating system. The monogamy rules are as follows. Choosy females mate with conspicuous males, one to each male, so long as any remain. If there are more choosy females than there are conspicuous males, the remaining choosy females mate with cryptic males. Undiscriminating females mate at random among those males remaining when the choosy females have made their choice. The products of the four possible matings are thus weighted as follows:

| if $r' > p'$ then ($C \heartsuit$ | $\times A \delta$) has weight p' |
|--|-------------------------------------|
| $(C \ \ \times \ a \ \)$ | (r' - p') |
| $(c \stackrel{\circ}{\scriptstyle{\leftarrow}} \times A \stackrel{\circ}{\scriptstyle{\leftarrow}})$ | 0 |
| $(c \ \heartsuit \ \times \ a \ \eth)$ | (1 - r') |
| if $r' \leq p'$ then $(C \not\subseteq$ | $\times A \delta$) has weight a |

 $\begin{array}{ll} (C \heartsuit \times a \eth) & 0 \\ (c \heartsuit \times A \eth) & (p' - r') \\ (c \heartsuit \times a \circlearrowright) & (1 - p') \end{array}$

Here, r' is the frequency of the C gene

among females after selection, and p' is the frequency of the A gene among males after selection. These rules guarantee that there is no automatic mating advantage; an increase in the frequency of the C gene does not of itself tend to cause an increase in the frequency of the A gene. Fisherian sexual selection cannot operate in a system of this sort, since all males are eventually mated to a single female, given that the sex ratio is equal. No advantage can accrue to conspicuous males, other than through the increased mean fitness of their offspring, and therefore any increase in the frequency of conspicuous males can be unequivocally attributed to the handicap principle.

The polygamy mating rules permit only three of the four possible matings. Choosy females mate only with conspicuous males, whilst undiscriminating females mate with cryptic and with conspicuous males in proportion to their frequencies in the population. The four possible matings thus receive the weights:

| $(C \ \ \times A \ \ \delta)$ | has weight r' |
|--|------------------|
| $(C \stackrel{\bigcirc}{+} \times a \stackrel{\circ}{\circ})$ | 0 |
| $(c \stackrel{\bigcirc}{} \times A \stackrel{\circ}{\circ})$ | p'(1 - r') |
| $(c \stackrel{\bigcirc}{\rightarrow} \times a \stackrel{\frown}{\circ})$ | (1 - p')(1 - r') |

The polygamy rules were used by Maynard Smith (1976) in his critique of the handicap principle, and by O'Donald (1962, 1963, 1967) in his extensive analytical and numerical investigations of Fisherian sexual selection. Conspicuous males have a mating advantage over cryptic males because they are guaranteed all the matings of choosy females, while at the same time receiving a proportion of the matings of undiscriminating females; any increase in the frequency of the C gene will thereby automatically cause an increase in the frequency of the A gene. An increase in the frequency of conspicuous males may be due to Fisherian sexual selection, or to sexual selection under the handicap principle, or to both acting jointly.

The simulation then proceeds as follows. The input variables IfS and FS are given, and an initial population of zygotes constructed in linkage equilibrium. The values of r' and p' are found by operating on the zygote genotype frequencies with survival rates (the 'fitnesses' in FS). For each of the four possible matings, the gene frequencies after selection at the three loci amongst the parents are computed, and are used to generate offspring genotypes in linkage equilibrium. The frequencies of these offspring genotypes are then weighted according to the mating rules adopted for the simulation, using the values of r'and p' obtained previously, and summed for all four possible matings. This procedure results in the calculation of genotype frequencies in the next zygote generation; these are used as a source of output data, and recycled to compute genotype frequencies in the following generation. In this manner, the genotype frequencies can be followed for as many generations as is desired. A listing of the program is available from the author on request; aside from the monogamy mating rules, it is identical with that used by Maynard Smith (1976).

Results.-I shall describe first the results of simulations governed by the monogamy mating rules, by which means Fisherian sexual selection is excluded. If the simulation is initiated with $IfS = \{q_0\}$ $p_0 = r_0$ and $FS = \{\alpha > 0, \beta > 0, \epsilon >$ 0}, then after the first generation of selection we invariably find that: (i) $p_1 < p_0$; (ii) $q_1 > q_0$; (iii) $r_1 = r_0$; (iv) $D_{BC} \equiv$ $(\boldsymbol{P}_{BC}\boldsymbol{P}_{bc} - \boldsymbol{P}_{Bc}\boldsymbol{P}_{bC}) > 0$. No selection has yet been applied at the C locus, where the alleles therefore do not change in frequency. The *B* gene is unconditionally favored and therefore increases in frequency. The linkage disequilibrium parameter D_{BC} , which measures the excess or deficiency of coupling genotypes at the B and C loci, becomes positive for the following reason. C females tend to mate with A males, while the B allele is more frequent amongst the A individuals than among the male population as a whole, if $q_A > q_a$. In this manner, B and C tend to become associated, so that the population contains an excess of coupling genotypes at these two loci. The effect is a small one, the

greatest observed value of D_{BC} not exceeding 10⁻⁴. If we put $\beta < 0$, then $q_A < q_a$ and $D_{BC} < 0$, indicating the expected excess of repulsion genotypes. The failure to find any region of parameter space in which $p_1 > p_0$ is at first sight disturbing, since the analysis clearly indicates that such a region should exist. This result, however, follows from the design of the program.

The inequality (3) is not equivalent to a condition that the initial frequency of the A gene should be exceeded in the first or in any subsequent generation of selection, under the conditions of the simulation. We begin the process in generation j = 0 with a population of zygotes in linkage equilibrium. Natural selection then acts on these zygotes through differential survival, as the result of which the frequency of the A gene must decrease if conspicuous males run some additional risk of mortality. Since the A gene is under-represented in the adults, it will also be under-represented in the next zygote generation (j = 1), assuming as we have that the C gene confers no automatic mating advantage. Therefore the frequency of the A gene among zygotes will always decline during the first generation of selection, so that $p_1 < p_0$. However, if condition (3) is satisfied, the zygotes from the matings of Amales will be on average more fit than those from the matings of a males, and the A gene will now tend to increase in frequency. Since it is now over-represented in the adult population, it will also be over-represented in the next zygote generation (j = 2), so that $p_2 > p_1$. This argument suggests that, provided condition (3) is satisfied, a population which is initially in linkage equilibrium will experience a decrease in the frequency of the A gene between the 0th and 1st zygote generations under natural selection and an increase between the 1st and 2nd zygote generations under sexual selection; if the simulation were to begin with a population of adults, then the A gene could increase between the 0th and 1st zygote generations. However, this process will be affected by the generation of linkage disequilibrium and by the further constraints on fitness predicated by the mating rules, and must therefore be investigated through simulation.

The analysis suggests that a necessary condition for sexual selection is $\epsilon \gg \beta$, and in the following experiments $FS = \{0.01, 0.001, 0.1\}$ was adopted. Condition (3) can be rearranged to read:

$$(q_A - q_a) > \beta/2\epsilon. \tag{4}$$

From equations (1) it follows that:

$$(q_A - q_a) = \frac{\beta \epsilon q (1 - q)}{(\alpha + q \epsilon)(\alpha + \beta + q \epsilon)}.$$

Substitution into (4) yields the quadratic inequality:

$$Aq^2 + Bq + C > 0$$

where $A = -3\epsilon^2$; $B = \epsilon(2\epsilon - 2\alpha - \beta)$; $C = -\alpha(\alpha + \beta)$. This has two real positive roots for the FS given: $q \approx 0.016$ and 0.58. Sexual selection should be sufficiently powerful to cause an increase in the frequency of the A gene if q_0 lies between these limits. The greatest effect of sexual selection will occur at the value \hat{q} which is defined by:

$$\frac{\partial (q_A - q_a)}{\partial q} \Big|_{q} = \hat{q} = 0$$

Performing the differentiation yields:

$$X\hat{q}^2 + Y\hat{q} + Z = 0$$

where $X = \epsilon(2\alpha + \beta + \epsilon)$; $Y = 2\alpha(\alpha + \epsilon)$ β ; $Z = -\alpha(\alpha + \beta)$. For the FS given, $\hat{q} \approx 0.087$. Thus, the analytical treatment suggests that the A gene should be favorably selected under the given FS in the range 0.016 $\lesssim q_0 \lesssim$ 0.58 and that this effect should be strongest when $q_0 = 0.087$. Figure 1 shows the results of simulation. The point of maximum Δp is correctly predicted by the analysis, but the entire range over which $\Delta p > 0$ is greatly overstated. Thus, the conditions shown by simulation are more stringent than those suggested by analysis; this is because the FS implies such an enormous selective advantage for the B gene that if it begins at a frequency of more than 10% or so it is almost fixed by the time that the second



FIG. 1. Sexual selection under the handicap principle in a monogamous population. Ordinate = change in frequency × 10⁵. Solid line = Δp ; broken line = Δr . Δp becomes positive over part of the range of q_0 . An arrow indicates the value of q_0 at which the maximum of Δp is predicted by analysis. Input variables: $IfS = \{0.001, q_0, 0.001\}$; $FS = \{0.01, 0.001, 0.1\}$. Change in frequency of A is measured between 1st and 2nd zygote generations, so that $\Delta p = p(j = 2) - p(j = 1)$.

zygote generation is formed. I emphasize that although the handicap principle is indeed capable of generating increases in the frequency of conspicuousness, these increases are both transient and very small, and are unlikely to constitute an evolutionary process of general importance.

Obviously, sexual selection will occur more easily if the B gene takes longer to pass through the population, since it will then segregate in the appropriate range of frequencies for a longer period of time. The range of q_0 over which $\Delta p > 0$ in Figure 1 might therefore be increased if sexual selection is effective for $\epsilon < 0.1$. To investigate this possibility, experiments were performed with $IfS = \{0.001, 0$ 0.001 and $FS = \{0.01, 0.001, \epsilon\}$; the results are shown in Figure 2. Since the fitnesses are intended as rates of survival, ϵ cannot exceed 1 - (0.01 + 0.001) =0.989. For values which approach this limit, $\Delta p > 0$ in the second zygote gen-



FIG. 2. Sexual selection under the handicap principle in a monogamous population. Δp varies with differing strengths of the effect at the *b* locus. Input variables: *IfS* = {0.001, 0.001, 0.001}; *FS* = {0.01, 0.001, ϵ }. *j* = generation.

eration, which is consistent with the analytical theory. For progressively smaller values of ϵ , the following pattern of events emerges. There is an excess of coupling genotypes $(D_{BC} > 0)$ in the first zygote generation, as expected, and this excess increases as long as $(q_A - q_a)$ increases. Eventually, q becomes so large that (q_A) $(-q_a)$ diminishes in time, and as it does so the excess of coupling genotypes also diminishes, finally leading to $D_{BC} = 0$ when q = 1. As $(q_A - q_a)$ and D_{BC} reach their maximum values, Δp becomes positive, and these positive values persist for between one and three generations. As ϵ decreases, Δq decreases and the maxima of $(q_A - q_a)$ and D_{BC} are shifted backwards in time, so that the region of increase in the frequency of the A gene occurs in successively later generations. Finally, when

 $\epsilon \leq 1/_{30}$, sexual selection becomes entirely ineffectual.

This result indicates the central dilemma of sexual selection under the handicap principle. Sexual selection occurs rather easily when ϵ is very large relative to β , but this is a very short-lived process because it implies a very rapid transit of the B gene through the population. If ϵ is decreased in order to slow down the passage of the B gene, then we quickly enter a region where ϵ is insufficiently large to procure any increase in the frequency of the A gene. The fixation of the B gene can also be delayed by making α large relative to ϵ , whilst keeping the ratio ϵ/β large, as for the example in $FS = \{0.5, 0.001,$ 0.1]. However, this strategem has the additional effect of reducing the rate of accumulation of B genes by A males and for this reason fails to provide a solution to the dilemma.

Although it is not difficult to confirm that transient increases in the frequency of the A gene may occur, in no instance was it found that the A gene could exceed its initial frequency; that is, although $p_i > p_i$ p_{i-1} (for some j) for some IfS and FS, $p_i < p_j$ p_0 (for all j > 0) for all IfS and FS investigated. It is not inconceivable that there may exist regions of parameter space in which $p_i > p_0$, but these must be very small, if they exist at all. I conclude that a sustained increase in the frequency of conspicuous males will occur only if different 'fitness genes' (e.g. those at the Blocus) are segregating within a narrow band of frequencies for long periods of time. This implies not only very powerful directional selection in each generation, but also a continuous supply of mutants with very large fitness differentials on which such selection can act.

These results are very sensitive to the two given of the model other than *IfS* and *FS*: the mating rules and the initial linkage equilibrium. Consider first the effect of substituting the polygamy for the monogamy mating rules. If we have *IfS* = $\{0.001, 0.001, 0.1\}$ and *FS* = $\{0.001, 0.001, 0.1\}$, then sexual selection is very powerful, the *A* gene being fixed and the



FIG. 3. The handicap principle in a polygamous population. Pure Fisherian sexual selection is modelled by $\epsilon = 0$. The handicap principle results in greater equilibrium frequency of A. Input variables: IfS = {0.001, 0.001, 0.001}; FS = {0.01, 0.001, ϵ }.

C gene reaching an equilibrium frequency of about 0.39. However, the high initial frequency of the C gene relative to that of the A gene gives so great a mating advantage to the conspicuous males that a very similar result would have followed from straightforward Fisherian sexual selection alone. This has been pointed out previously by Maynard Smith (1976). However, in his simulation he set u = 1 (see description of his fitness model, above), which is equivalent to the assumption that fitness effects are multiplicative; his statement that Fisherian selection works best if there is no handicap is not necessarily correct if fitness effects are not multiplicative. We may model a pure Fisherian process by putting $\epsilon = 0$; the *B* gene is now neutral, the genotypic fitnesses are obtained multiplicatively, and conspicuous males can accumulate no excess of potentially advantageous genes at other loci. If we put $IfS = \{0.001, 0.001$ 0.001, the mating advantage of conspicuous males is sufficiently great to procure an increase in the frequency of the A gene, which rises to an equilibrium frequency of between 0.09 and 0.10. By putting $\epsilon > 0$, we can study the effect of the handicap principle on this process. Figure 3 shows that as ϵ increases both the rate of increase and the equilibrium frequency of the A



FIG. 4. The handicap principle in a polygamous population. The handicap principle may cause an increase in the frequency of *A*, which decreases under a pure Fisherian process. Input variables: *IfS* = $\{0.001, 0.001, 5 \times 10^{-5}\}$; *FS* = $\{0.01, 0.001, \epsilon\}$.

gene increase; for $\epsilon = 0.1$, the A gene eventually comes close to fixation. This result raises the possibility that there may be parameter values under the polygamy mating rules for which pure Fisherian sexual selection is ineffective but for which a combination of the Fisher effect and the handicap principle may procure an increase in the frequency of conspicuousness. Specifically, as the initial frequency of the C gene is reduced whilst the initial frequency of the A gene is kept constant, there exists a point below which natural selection of given intensity will always be effective in reducing the frequency of the A gene. For $FS = \{0.01, 0.001, 0\}$ this point occurs in $10^{-4} > r_0 > 9 \times 10^{-5}$ (clearly, at $r_0 = 10^{-4}$ the 10% disadvantage of the A gene under natural selection will be just exceeded by the slightly greater than 10% advantage under sexual selection given by the polygamy mating



FIG. 5. Initial stages of sexual selection in a polygamous population. The handicap principle may enable a Fisherian process to become established. (Top) Change in frequency of A and C. (Middle) Magnitude of Δp . (Bottom) Increase in linkage disequilibrium between A and C loci: $D_{AC} = P_{AC}P_{aC} - P_{AC}P_{aC}$. Input variables: $IfS = \{0.001, 0.001, 5 \times 10^{-5}\}$; $FS = \{0.01, 0.001, \epsilon\}$.

rules). Thus, if we set $r_0 = 5 \times 10^{-5}$ and $\epsilon = 0$, the A gene declines from an initial frequency of 10^{-3} to some lower equilibrium value (in this case, about 3×10^{-4}). Pursuing the argument, we expect this lower equilibrium value to increase with ϵ , and speculate that for some ϵ it may exceed the initial frequency of the A gene. Figure 4 shows that this is indeed the case: for $\epsilon = 0.1$ the A gene increases to an eventual equilibrium at >1%. Two rather different processes are involved in this event. First, the accumulation of B genes by A males lessens the disadvantage of their A-bearing male progeny under natural selection, and secondly, the passage of the B gene through the population strengthens the coupling between A and



FIG. 6. Initial stages of sexual selection in a polygamous population. The strength of the *B* locus necessary for the handicap principle to enable a Fisherian process to become established. Input variables: $IfS = \{0.001, 0.001, 5 \times 10^{-5}\}; FS = \{0.01, 0.001, \epsilon\}.$

C and thereby increases the mating advantage of A males under Fisherian sexual selection. Both effects are small, but in combination they are sufficiently powerful to reverse the inequality between natural and sexual selection. The initial stages in the evolution of the population are illustrated in Figure 5. While the B gene is rare, $\Delta p < 0$; but by the third generation q has become large enough to make q_A – q_a) sufficiently large to fuel the handicap principle, so that $\Delta p > 0$. As $q \rightarrow 1$ the power of the handicap principle wanes and eventually disappears, but by this time r and D_{AC} (the linkage disequilibrium between the A and C loci, D_{AC} = $P_{AC}P_{ac} - P_{AC}P_{aC}$ have become sufficiently large for the A gene to continue to be favorably selected by a pure Fisherian process.

It is clear that for any given value of ϵ there will be a value of r_0 below which even the combination of Fisherian sexual selection and the handicap principle will fail: for $\epsilon = 0.1$ this value is nearly 10^{-5} . Thus, the effect of the handicap principle in this experiment is to extend the range of frequencies of the *C* gene over which sexual selection is effective downward by about one order of magnitude. Similarly, we might ask what strength of effect at the *B* locus is required in order to overcome the overall disadvantage of *A* males. Figure 6 illustrates the initial dynamics of selection for different values of ϵ ; if $\epsilon >$ 0.01, then the eventual equilibrium frequency of the *A* allele lies above its initial frequency. The eventual frequency reached by the *A* gene when $\epsilon = 0.1$ exceeds 1%.

Maynard Smith (1976) did not find that the handicap principle would work when choosy females are rare relative to conspicuous males, but in the only experiment that he reports he sets $r_0 = p_0/500$. Since the positive result reported above held only between $r_0 = p_0/10$ and $r_0 = p_0/100$, it seems likely that his experiment merely confirms that there is a region of $r_0 \ll p_0$ in which even the combination of the handicap principle and the Fisher effect cannot induce effective sexual selection.

The analytical results suggest that sexual selection under the handicap principle will be easier if the population initially includes an excess of coupling genotypes, and this is readily confirmed by simulation. It appears that if the initial population consists exclusively of coupling genotypes (BC and bc), the A gene will increase even under the monogamy mating rules if $\epsilon > \beta$. The initial rise in frequency is not surprising. Since C is coupled with B, C will increase in frequency as B increases under natural selection. At the same time, C will come into coupling with A because of the mating preference of C females for A males, and the resultant coupling of A with B causes an increase in the frequency of A, provided that the average effect of the B gene (measured by ϵ) exceeds that of the A gene (measured by β). This is merely an example of the 'hitch-hiker' effect. More surprisingly, although the A gene will subsequently decline in most cases, there are parameter values for which A continues to increase until B is fixed. This is because a population consisting only of ABC and abc zygotes is equivalent to two non-interbreeding subpopulations, and A is certain to be fixed if the fitness of AB individuals exceeds that of *ab* individuals, which requires $\epsilon > \beta$. Similar results follow from less extreme initial linkage disequilibria.

DISCUSSION

The analysis and the simulations reported above demonstrate the formal possibility that the handicap principle can generate effective sexual selection. The conditions under which this possibility is realized can be scrutinized in order to throw light on the applicability of the handicap principle to real populations.

A non-multiplicative model of genotypic fitnesses is essential to the operation of the handicap principle, and the arguments presented above are all derived from a model of additive gene effects. The distinction between the additive and multiplicative ways of combining effects at different loci is as follows. The additive model implies that effects on fitness are mutually exclusive but not independent; the multiplicative model implies that effects in fitness are independent but are not mutually exclusive. Effects which are neither independent nor mutually exclusive are combined according to models intermediate between the additive and multiplicative. For example, Ricklefs (1977) uses an additive model of mortality when 'reproductive' and 'non-reproductive' risks are run concurrently, but he uses a multiplicative one when the exposure to these two different sources of mortality is separated in time. If, in the present example, effects at the A and B loci were completely separated in time, then the genes at the B locus would be neutral when selection was occurring at the A locus, and A males could accumulate no excess of B genes. Thus, a necessary condition for the handicap principle to operate at all is that the effects on fitness at the A and B loci should not be perfectly independent. This is not unreasonable; one would expect it to be satisfied, for instance, if the character controlled by genes at the B locus were expressed during the breeding season as well as outside it. J. Maynard Smith (pers. comm.) has pointed out to me that a fitness scheme in which conspicuousness has a much more deleterious effect on b

males than on B males may be more favorable to the operation of the handicap principle than the additive scheme used above. However, there seems to be little biological justification for this arrangement.

Under the monogamy mating rules, a transient increase of the A gene may occur if $FS = \{ \alpha \leq \epsilon \gg \beta \}$. The baseline fitness cannot be large relative to the effect at the *B* locus, since $\partial (q_A - q_a)/\partial \alpha < 0$, so that the accumulation of B genes by Amales is retarded by large α . The *B* locus must therefore represent a major determinant of fitness. At the same time, the effect at the B locus must be very large relative to the sex-limited effect at the A locus; since the parameter space has four dimensions it is not possible to summarize this condition with a single figure, but increase in the frequency of the A gene for two or more generations at a rate of 1% or more per generation was observed only when $\epsilon \gtrsim 100\beta$. This implies that β must be small relative to α , that is, that the effect of conspicuousness on survival, in the absence of the B gene, must not be very great. When these conditions on the fitness set are satisfied, the A gene increases in frequency under sexual selection in the range of q for which $(q_A - q_a)$ is sufficiently large. This range is not, however, very extensive (see Fig. 1). Moreover, the condition $FS = \{ \alpha \leq \epsilon \gg \beta \}$ implies a large unconditional advantage for the Bgene, which passes rapidly through the population under natural selection, persisting in the range throughout which sexual selection is effective for only one or a very few generations. The rate of passage of the B gene can be slowed only by increasing α or by diminishing ϵ or both, and these operations make effective sexual selection unlikely because of the conditions on the fitness set given above.

A sustained increase in the frequency of the A gene was not observed, and is probably impossible under the conditions of the simulation, given the monogamy mating rules and initial linkage equilibrium. It could be achieved only if the conditions on the fitness set are satisfied, while at the same time a succession of genes equivalent in effect to the B gene sweep through the population; the process envisaged could either be the rapid fixation of genes at different loci, or the alternation of selection coefficients at a single locus. Such extremely powerful, prolonged directional selection does not seem very plausible.

These stringent conditions on the operation of the handicap principle are greatly relaxed if the B and C genes are initially in coupling. However, although it is perfectly possible that they should happen to occur on the same chromosome, the process is something of a *deus ex machina*, and the handicap principle adds little to the 'hitch-hiker' effect that would occur in any case.

I conclude that the handicap principle is of dubious value in explaining the evolution of conspicuous secondary sexual characters in strictly monogamous populations. There is, however, no other known mechanism by which sexual selection can achieve this result (although the operation of sexual selection on other types of character, for example the date of breeding, in monogamous populations is well known; see Darwin, 1874, and O'Donald, 1972 and 1973). If the sex ratio amongst adults is equal, then each male will find a mating partner, and sexual selection can act only through the quality of offspring, rather than through the number of females inseminated. If conspicuous males are indeed at a disadvantage with respect to natural selection (a conclusion disputed by many authors, and notably by Thayer, 1909; and Cott, 1957), then their occurrence in monogamous species cannot be due to natural selection but at the same time receives no adequate explanation through Darwinian or Fisherian sexual selection. Moreover, this is not a negligible phenomenon. About a quarter of the monogamous North American passerines listed by Verner and Willson (1969) are sexually dimorphic, including such striking and well-studied species as the cardinal, *Richmondena cardinalis* (male bright red, female predominantly yellow-brown), and the rufous-sided towhee, Pipilo erythophthalmus (male with head and entire upper body black, female dusky brown).

Polygamous mating systems are thought to generate powerful sexual selection, because there can be considerable variation in the number of females inseminated by different males. This idea is originally due to Darwin (1874), who regarded certain male characters as necessarily stimulating the female to mate, without explaining how this necessity arose or could be maintained. The modern version of Darwin's theory is due to Fisher (1930), who pointed out that the female's ability to choose, as well as the male's predisposition to be chosen, will be liable to selection. The male character in question is supposed to be favored by natural selection when it first appears in the population. Because those males which express the character are the more fit, females which prefer to mate with them will be favored by virtue of the increased mean fitness of their offspring. The males are now doubly favored, since the increment in fitness caused by the mating preference is added to their prior advantage under natural selection. The male character and the female preference spread through the population; and as they do so, any alternative genes or modifiers which exaggerate the expression of the male character or increase the powers of discrimination of the female are also favored. The exaggeration of the male character eventually renders it harmful to its bearer, so that it begins to be exposed to unfavorable natural selection. But by this time the reproductive advantage that the character enjoys by virtue of the female mating preference is sufficiently large to overpower its disadvantage under natural selection. Indeed, the overall advantage of the character tends to be directly frequency-dependent, since the reproductive advantage increases with the frequency of choosy females, which in turn increases with the frequency of conspicuous males. The result is a 'run-away' process by which the genes for conspicuousness are rapidly fixed, or reach a stable equilibrium at which the effects of sexual and natural selection cancel one another.

This mode of sexual selection is very powerful if the polygamy mating rules described above are given, since each Amale recieves on average (r'/p') + (1 - r')females, whereas each a male receives on average only (1 - r') females. Countervailing natural selection will eliminate the A males only if it is sufficiently intense to overcome the reproductive advantage generated by an additional (r'/p') matings. However, certain objections to the applicability of Fisher's theory to natural populations should be mentioned. The first is that if the current disadvantage of the male character with respect to natural selection is granted, it is still required that on its first appearance, in a less exaggerated form, the character should actually have reduced the risk of mortality. Although not inconceivable, it is not clear that this should be true for the majority of 'conspicuous' characters. Secondly, the polygamy mating rules themselves may not be ecologically plausible.

Consider the initial stages of selection, when the male character is still favored by natural selection. If males bearing the character are to receive a mating advantage (in terms of the number of broods they sire) in a population where the sexes are equally frequent, then they must on average mate with more than one female, while males not bearing the character attract only one female, or do not succeed in mating at all. This implies that the increased fitness of offspring more than compensates a choosy female for any disadvantages incurred by sharing a single male with other females. This is not implausible if there is no lasting pair-bond, and if the female alone, or neither sex, is responsible for parental care of the young; but if the male alone or the male and female together care for the young, the heritable increase in the mean fitness of their offspring must exceed the effect of any impairment of their offspring's survival and growth caused by diminished per capita parental care. Since the effect of polygyny on female fitness may be rather severe (e.g. Downhower and Armitage, 1971),

this seems to require an unrealistically high heritability of fitness.

This objection, of course, applies equally to Fisherian sexual selection and to the handicap principle: what is under attack is not the reality of either of these processes, but, rather, the reasonableness of the polygamy mating rules in species where the male is involved in parental care. It seems plausible to suppose that in real populations with pair-bonds during the breeding season a certain proportion of choosy females will prefer to mate with a cryptic male, rather than being the second, third or n-th mate of a conspicuous male. If this proportion is zero, the polygamy mating rules hold; if it is unity, the monogamy rules will be appropriate. As the proportion passes from zero through intermediate values to unity, we can expect two things to happen: first, the overall strength of sexual selection will decline; and secondly, the importance of the handicap principle relative to the Fisher effect will increase, until at the limit only the handicap principle is capable of producing any increase in the frequency of genes for conspicuousness. This argument implies the existence of a range of mating rules within which a pure Fisherian process will be impotent to prevent the decline of genes for conspicuousness, while the combination of Fisherian selection and the handicap principle procures effective sexual selection.

Maynard Smith (1976) has pointed out that the handicap principle will be effective in the absence of any Fisher effect, if there is no additive genetic variance for male conspicuousness. The B genes accumulated by conspicuous males are then passed on to the progeny of choosy females, without any systematic disadvantage being conferred by inheritance of the handicap. Although this possibility is dismissed as being special pleading, it is not implausible, except that the characters of interest in real populations are probably nearly always inherited. However, consider the course of evolution in a population where there is initially no additive genetic component to the variation of conspicuousness in males. The male character is obviously indifferent to selection and will remain at the same frequency throughout. Female choosiness, however, is favored, and the genes responsible will increase in frequency. Suppose that genic variance for conspicuousness now arises, for example by mutation $a \rightarrow A$ at the A locus. If the population is strictly monogamous, then conspicuousness is not favored except under the stringent conditions described above. But if the population is polygamous, then the relatively high frequency of choosy females will induce Fisherian sexual selection, and the A gene will increase in frequency. Thus, one major difficulty of Fisher's theory of sexual selection-the prior segregation of genes for female choosiness at appreciable frequency—can be removed by supposing that genes for male conspicuousness mimic previously non-heritable characters, without the need to suppose that the male character is advantageous with respect to natural selection when expressed in a less exaggerated fashion.

The interaction of the handicap principle with Fisherian sexual selection may be its most important, or only, contribution to the evolution of sexual dimorphism. The augmenting of Fisherian selection by the handicap principle might be critical if female choosiness were rare relative to male conspicuousness, if the mating rules of the population tended to monogamy, or if there were originally no additive genetic variance for conspicuousness. For example, the handicap principle might play a part in the evolution of male secondary sexual characters in populations of birds which although primarily monogamous exhibited occasional polygyny. The critical evaluation of these possibilities, and their relevance to the field situation, awaits the exploration of models with diploid heredity and realistic mating rules.

SUMMARY

Zahavi (1975) has suggested that females may prefer to mate with conspicuous males because these males, having survived despite their handicap, must be extraordinarily fit in other respects. This paper examines the validity of the handicap principle by studying the behavior of a simple haploid model. Three loci are involved: alternative genes at the A locus determine male conspicuousness (A---conspicuous; a—cryptic); those at the B locus determine some unrelated aspect of fitness (B—more fit; b—less fit); those at the Clocus determine the mating preference of females (C-choosy, prefer to mate with A males; *c*—undiscriminating). If the effects on fitness at the A locus (measured by β) and at the B locus (measured by ϵ) are multiplicative, the handicap principle cannot work. If these effects are additive, sexual selection in a monogamous population follows if $2\epsilon(q_A - q_a) > \beta$, where q_A and q_a are the frequencies of the B gene after selection in conspicuous and cryptic males respectively. However, although this condition leads to small and transient increases in the frequency of the A gene, sustained increase is contingent not only on very powerful and prolonged direction selection, but also on a continuous supply of mutants providing very large fitness differentials on which this selection can act. These conditions are so restrictive that the handicap principle acting alone is unlikely to be a significant evolutionary force. In polygamous populations, the handicap principle can be thought of as modulating the dynamics of Fisherian sexual selection, leading in general to higher equilibrium frequencies of the A gene. In particular, the combination of the handicap principle and Fisherian sexual selection can generate nontrivial increases in the frequency of male conspicuousness in circumstances where a pure Fisherian process cannot: this may happen if female choosiness is initially rare relative to male conspicuousness; if the mating system is intermediate between monogamy and polygyny; or if the A gene mimics

previously non-heritable variation. It is concluded that any importance of the handicap principle in real populations is likely to reside in its catalytic effect on Fisherian sexual selection.

Acknowledgments

This paper has grown out of a correspondence with J. Maynard Smith, and could not have been written without his advice. I am also indebted for comments and criticisms to P. Handford, U. Motro, P. O'Donald, R. Shand, J. Turner and A. Zahavi.

LITERATURE CITED

- COTT, H. B. 1957. Adaptive Colouration in Animals. Methuen, London.
- DARWIN, C. R. 1874. The Descent of Man and Selection in Relation to Sex. 2nd Edn. John Murray, London.
- DOWNHOWER, J. F., AND K. B. ARMITAGE. 1971. The yellow-bellied marmot and the evolution of polygamy. Amer. Natur. 105:355–370.
- FISHER, R. A. 1930. The Genetical Theory of Natural Selection. Clarendon Press, Oxford.
- MAYNARD SMITH, J. 1976. Sexual selection and the handicap principle. J. Theoret. Biol. 57:239–242.
- O'DONALD, P. 1962. The theory of sexual selection. Heredity 17:541-552.
- ——. 1963. Sexual selection for dominant and recessive genes. Heredity 18:451–457.
- ------. 1967. A general model of natural and sexual selection. Heredity 22:499–518.
- ------. 1972. Natural selection of reproductive rates and breeding times and its effect on sexual selection. Amer. Natur. 106:368–379.
- ——. 1973. Frequency-dependent sexual selection as a result of variations in fitness at breeding time. Heredity 30:351–368.
- RICKLEFS, R. E. 1977. On the evolution of reproductive strategies in birds: reproductive effort. Amer. Natur. 111:453-478.
- THAYER, G. H. 1909. Concealing Colouration in the Animal Kingdom. MacMillan, New York.
- VERNER, J., AND M. F. WILLSON. 1969. Mating systems, sexual dimorphism, and the role of male North American passerine birds in the nesting cycle. Ornithological Monographs, No. 9. American Ornithologists' Union.
- ZAHAVI, A. 1975. Mate selection—a selection for a handicap. J. Theoret. Biol. 53:205–214.

http://www.jstor.org

LINKED CITATIONS

- Page 1 of 1 -

STOR

You have printed the following article:

The Handicap Principle in Sexual Selection Graham Bell *Evolution*, Vol. 32, No. 4. (Dec., 1978), pp. 872-885. Stable URL: http://links.jstor.org/sici?sici=0014-3820%28197812%2932%3A4%3C872%3ATHPISS%3E2.0.CO%3B2-H

This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.

Literature Cited

The Yellow-Bellied Marmot and the Evolution of Polygamy

Jerry F. Downhower; Kenneth B. Armitage *The American Naturalist*, Vol. 105, No. 944. (Jul. - Aug., 1971), pp. 355-370. Stable URL: http://links.jstor.org/sici?sici=0003-0147%28197107%2F08%29105%3A944%3C355%3ATYMATE%3E2.0.CO%3B2-Z

Natural Selection of Reproductive Rates and Breeding Times and Its Effect on Sexual Selection

Peter O'Donald *The American Naturalist*, Vol. 106, No. 949. (May - Jun., 1972), pp. 368-379. Stable URL: http://links.jstor.org/sici?sici=0003-0147%28197205%2F06%29106%3A949%3C368%3ANSORRA%3E2.0.CO%3B2-D

On the Evolution of Reproductive Strategies in Birds: Reproductive Effort

Robert E. Ricklefs *The American Naturalist*, Vol. 111, No. 979. (May - Jun., 1977), pp. 453-478. Stable URL: http://links.jstor.org/sici?sici=0003-0147%28197705%2F06%29111%3A979%3C453%3AOTEORS%3E2.0.CO%3B2-X