SCAN SAMPLING AND WATERFOWL ACTIVITY BUDGET STUDIES: DESIGN AND ANALYSIS CONSIDERATIONS

by

GRAHAM HEPWORTH\(^1,2\) and ANDREW J. HAMILTON\(^3,4\)

\(^1\) Department of Mathematics and Statistics, The University of Melbourne, Vic. 3010, Australia; \(^3\) Applied Ornithology Group, Johnstone Centre, School of Environmental and Information Sciences, Charles Sturt University, P.O. Box 789, Albury, NSW 2640, Australia, and Institute for Horticultural Development, Private Bag 15, Scoresby Business Centre, Vic. 3176, Australia)

(Acc. II-X-2001)

Summary

Scan sampling is a common technique used to quantify the activities of animals, including waterfowl. When large numbers are present, it is often impractical to record the activities of each individual, and sub-sampling is employed. We present a method for the design and analysis of scan sampling studies involving sub-sampling, based on an actual study of waterfowl activity on a waste stabilisation pond. The design we propose avoids subjectivity in the selection of individuals, is truly random rather than haphazard, and is adaptable to other situations. It allows for the population on each sampling occasion to be divided into separate strata with samples taken from each. The method of analysis addresses the statistical issues arising from such designs. In estimating the proportion of individuals engaged in an activity, it uses the information from observations in which some of the individuals are counted but their activities not recorded, and estimates the sampling variance introduced by sub-sampling. In comparing the mean proportions for different levels of a factor, such as time of day, the method uses restricted maximum likelihood (REML), an algorithm which can account for the sampling variance as well as for missing data. It is anticipated that the methods described

---

\(^2\) Corresponding author; e-mail address: g.hepworth@ms.unimelb.edu.au

\(^4\) The authors thank Ms Sue Welham, Rothamsted Experimental Station, U.K., for developing Genstat code to incorporate sampling variance into the REML analysis, and Mr Wayne Robinson for comments on the experimental design in the early stages of the work. We are grateful to Dr Ray Watson and Dr Nam Ky Nguyen for constructive comments on the manuscript. We also acknowledge the support of the Johnstone Centre, Charles Sturt University, for covering the costs of transport to and from the field site.
in this paper will assist researchers using the scan sampling technique, or workers in any discipline where sampling characteristics are similar.

Introduction

The ‘activity budget study’ has been recognised as an important tool for understanding habitat use and niche separation (e.g. Rave & Baldassarre, 1989), and consequently it is an invaluable aid for managing waterfowl communities and habitats. Waterfowl activity budget studies can typically be grouped into either those that employ focal-animal sampling or those that use scan sampling. Focal-animal sampling involves observing an individual — normally chosen at random — for a known period of time (Altmann, 1974). Behaviour is usually recorded at pre-determined time points (Miller, 1976; Asplund, 1981; Tome, 1991; Rave & Baldassarre, 1989; Zicus & Hennes, 1993; Petrie & Rogers, 1997; Petrie & Petrie, 1998), although another approach has been to record to length of time the focal individual is engaged in a particular behaviour (e.g. Bergen et al., 1989; Józkowicz & Górskaklek, 1996). Scan sampling requires that the behaviour of individuals in the sample be recorded ‘instantaneously’ (Altmann, 1974). Altmann (1974) considered the advantages, disadvantages and appropriateness of both approaches. Whilst both methods have been used widely for waterfowl activity budget studies, this paper is concerned only with scan sampling.

Many waterfowl activity budget studies utilising scan sampling involve surveying the entire ‘local population’ at the time of sampling, e.g. all birds on a pond (Skead, 1977; Norman et al., 1979; O’Donoghue & O’Halloran, 1994; Adair et al., 1996). However, due to the large number of birds present, some studies have necessitated sub-sampling of a flock. Various methods have been employed to do this. A common approach has been to start at one end of a flock and observe individuals until a certain number is reached. Squires & Anderson (1997) scanned flocks of trumpeter swans from left to right, recording the activities of a maximum of 25 birds; the remaining birds were sampled in the next scan (scans made every ten minutes). Davis et al. (1989) scanned snow geese on land and water, beginning at one end and zigzagging across the flock until a set number were observed (600 or 300). Such an approach has the advantage of simplicity, but is disadvantaged by the selection of birds being non-random, and therefore not satisfying an assumption upon which statistical analysis is usually based. Starting at one
end and stopping at a certain number may result in birds in the middle of the flock having a smaller chance of being included in the sample than those at the end. Zigzagging across a flock is haphazard rather than random in the statistical sense, and is somewhat subjective. A method of sub-sampling which avoids these limitations is to divide the flock into separate sectors, randomly select at least one (preferably more) of the sectors, and observe all the birds within a visual field in the selected sectors. This method, which will be presented in this paper, allows the sub-samples to be spatially distributed across the flock rather than concentrated towards one end, and there is no subjectivity involved.

An approach with some similarities to this is to use random transects through a flock. Miller (1985) recorded the behaviour of all birds along a ‘line of sight in the middle of [the] field of view of a spotting telescope.’ This type of approach raises another simple but important question: how should sub-samples be chosen randomly? Miller (1985) described his transects as being chosen at random, but how randomness was achieved was not described.

The number of birds observed in the samples is another important factor in the success of an activity budget study, as it has a large bearing on the precision of estimates arising from the study, that is, on the variances associated with estimates of proportions and mean numbers. To some extent the precision is beyond the investigator’s control, but low numbers can be compensated for by sampling on more occasions. Even where the number of birds in a sample is low, the information in the sample should still be used in the analysis, though it needs to be weighted appropriately. Davis et al. (1989) omitted scans with fewer than 100 individuals from the analysis of their data; this may have simplified calculations, but at the expense of information, and ultimately precision.

In some instances the study site is too large to be sampled from one point (e.g. Campbell, 1978). A solution to this problem is to divide the site into non-overlapping sections that are observed separately. The data then need to be weighted according to the numbers of birds observed in the different sections, which adds a further level of complexity to that arising from sub-sampling at the section level (if sub-sampling has been used). Here we present an approach for situations where the study site is divided into sections, and sub-sampling is employed in some sections on some sampling occasions. We also consider statistical issues associated with data collected
in such a manner, in particular, the additional variance arising from sub-sampling.

The aim of this paper is to highlight specific issues associated with the design and analysis of waterfowl scan sampling studies — especially sub-sampling — and present a method for analysing data arising from such studies. This is primarily achieved through explaining how we approached the design and analysis of a study of waterfowl activity budgets on a waste stabilisation pond. That study is described in more detail in Hamilton et al. (in press). Ideas and methods put forward in this paper may be useful not only for waterfowl research, but for any study involving the scan sampling technique.

**Study site and species**

All observations were made at a waste stabilisation pond (Pond Nine) at the Western Treatment Plant (WTP), Werribee, Australia (38°00′S, 144°34′E). This pond is part of an extensive system of thirty sewage treatment ponds. It was chosen because it has previously been recognised as the most important site for waterfowl at the WTP (Elliget, 1980; Lane & Peak, 1990), and it is part of a wetland of international importance (Ramsar Convention Bureau, 1984).

The species studied were pink-eared duck (*Malacorhynchus membranaceus*), Australian shelduck (*Tadorna tadornoides*), black swan (*Cygnus atratus*), hardhead (*Aythya australis*), Australasian shoveler (*Anas rhynchotis*), chestnut teal (*A. castanea*), grey teal (*A. gibberifrons*) and Pacific black duck (*A. superciliosa*).

**Scan sampling**

Sampling was conducted at five evenly spaced times of day: sunrise; mid-morning; midday; mid-afternoon and sunset. It typically took approximately one hour to sample the entire pond, and hence sampling for each time of day commenced 30 minutes prior to the midpoint (*e.g.* 30 minutes before sunrise). Sampling was undertaken at approximately weekly intervals from 11/7/1998 to 20/6/1999 and then at approximately monthly intervals from 1/8/1999 to 9/8/2000.

Surveys were made from defined observation points on the embankment/road; vehicle access was available around the entire perimeter of the pond. All observations were made from within the car using a Leica® Televid 77 telescope (20-70× zoom magnification) mounted on the car window.
Aside from being more time efficient, counting from the car was preferable to using a tripod outside the car because it was felt that the birds would be less likely to be disturbed. The pond was too large to be sampled from one observation point. To facilitate sampling, the pond was divided into five sections. Note that the purpose of these sections was solely to aid surveying; they were not intended to represent a particular natural feature, and there was no intent to make comparisons between them. Nevertheless, one section — section C — was unique in that it contained an extensive stand of dead trees (*Melaleuca lanceolata*, approximately 3/4 of section) and nest boxes; all other sections consisted of open water and sometimes several small islands/exposed rocks (depending on water level). The order by which sections were sampled was randomised for each sampling date. This order was maintained throughout the day (*i.e.* for all five times of day).

Each of the five sections was scanned from left to right. One of the following activity classes was assigned to each individual of each species: feeding, resting, locomotion, comfort movement, or other.

A preliminary scan was made of each section. If this scan revealed more than approximately 100 individuals of a particular species, then this species ‘flock’ was sub-sampled. Otherwise, all birds of this species were observed. For one particular species, the pink-eared duck, such flocks were sometimes in the order of 2,000-5,000 individuals. A sub-sample consisted of a field of view, the positioning of which is described below. The number of birds in a sub-sample was dependent on two factors: distance of the flock from the observer, and ‘packing density’ of the flock. For most species, the mean number was about 50, the minimum around 10 and the maximum between 100 and 150.

Two sub-samples were taken on each occasion. The positioning of the field of view for a sub-sample was randomised according to the following procedure:

1. the field of view of the flock was divided into four approximately equal sized sectors (allocated a number, 1-4 from left to right);
2. two random numbers (from 1-4 inclusive) were chosen from a random number table;
3. the spotting telescope was then aligned with the approximate centre of each chosen sector and the appropriate activity data were recorded for each individual in the field of view.
For sub-sampled flocks, the total number of birds in the flock also needed to be estimated. This was achieved by counting birds in estimated groups of 10, 100 or 1,000 (depending on flock size).

**Estimation of proportions**

The aim of the sampling procedure described above was to obtain sound estimates of the numbers of individuals in each activity category, for each species on each sampling occasion. These numbers would then be used to estimate the proportion (or percentage) of individuals in each category, together with the associated sampling variance. The variance needed to be estimated for any estimates involving sub-sampling, because there was some uncertainty associated with such estimates, and it was appropriate to include this in the analysis of the data.

The sampling procedure had the following characteristics that had to be taken into account in the calculations:

(a) where sub-sampling was employed, not all individuals were categorised regarding activity;
(b) individuals within one of the sectors to be sub-sampled were more likely to be engaged in the same activity than individuals from different sectors, because of their proximity to one another;
(c) the five sections often had widely varying numbers of individuals.

For each species on each sampling occasion, the estimation of proportions proceeded as follows:

1. Where sub-sampling was used within a section, the numbers of individuals in each activity category were summed across the sub-samples. The resulting totals were used to estimate the proportion in each category, and these proportions were applied to the individuals not categorised, to give estimates of numbers in each category for the whole section. This addressed characteristic (a) above.
2. The sub-sampling of sectors employed in this study was a case of what in sampling theory is called ‘cluster sampling’. A feature of this method is that within a cluster, individuals are usually more homogeneous than would occur randomly (characteristic (b)). The variance of estimates arising from cluster samples is larger than that
obtained by a simple random sample from the entire population, and this should be reflected in the calculations. Cochran (1977, p. 66) gave an approximation for the estimated variance associated with an estimated proportion $p$ from a cluster sample, as follows:

$$v_c(p) = \frac{N - n}{N} \frac{\sum a_i^2 - 2p \sum a_i m_i + p^2 \sum m_i^2}{k(k - 1)\bar{m}^2}$$

(1)

where $N = \text{total population size}$; $n = \text{total sample size}$; $a_i = \text{number of individuals in the } i^\text{th} \text{ cluster in the category of interest, } i = 1, \ldots, k$ ($k = \text{number of clusters}$); $m_i = \text{number of individuals in the } i^\text{th} \text{ cluster}$; $\bar{m} = \text{mean of the } m_i$’s. This formula was applied to variance estimation from all sections involving sub-sampling; hence characteristic (b) was addressed. Another source of variation, not considered in this study, is in the estimation of $N$; however, even quite sizeable variation in $N$ results in only a small change to $v_c(p)$.

3. Following the estimation of proportions (step 1) and associated sampling variances (step 2) for each section, it remained to combine these into estimates relating to the entire pond. In sampling terminology, the sections would be called ‘strata’. The appropriate estimate for the proportion of individuals in each category across the whole pond is a weighted sum of the estimated stratum proportions, and it can be calculated according to the following formula (Cochran, 1977, p. 107):

$$p = \frac{\sum N_j p_j}{N}$$

(2)

where $N_j = \text{population size in the } j^\text{th} \text{ stratum, } j = 1, \ldots, 5$; $p_j = \text{estimated proportion in the } j^\text{th} \text{ stratum in the category of interest}$; $N = \text{total population size}$.

The application of this formula ensured that characteristic (c) was addressed.

The estimated sampling variance across the whole pond was the sum of the estimated variances from all sections; variances are additive for independent strata, and the sections could be validly assumed to be independent of each other. For sections not involving sub-sampling, the estimated sampling variance was zero, because the whole population present at that time was observed. The total sampling variance therefore came entirely from sections involving sub-sampling.
Statistical analysis

Following completion of the three steps described above, estimates of the proportion of individuals in each activity category, and the associated sampling variance, were available for each species on each sampling occasion. In accordance with most activity budget studies, the data were expressed as percentages. There were 235 sampling occasions, consisting of 47 sampling dates × 5 times of day. There were, however, many missing values for some species, corresponding to occasions when no individuals of the species were recorded.

The main purpose of the analysis was, for each species × activity category, to examine the effect of time of day on the percentage of individuals engaged in the particular activity. There was no interest here in formally comparing species or activities. The different sampling dates, which were at least a week apart, were assumed to be independent of each other in terms of the observed variables, and so each sampling date contained a ‘replicate’ of each of the times of day. In the terminology of experimental design, the sampling dates were taken as blocks, and the times of day as treatments assigned to plots.

The percentages calculated above were analysed using restricted maximum likelihood (REML) (Patterson & Thompson, 1971). The REML algorithm accounted for the sampling variance associated with many of the estimates, in addition to the large number of missing values. REML reduces to the standard analysis of variance (ANOVA) in simple balanced cases, i.e. where there are no missing values, and the only variance is that between, rather than within, observations.

The analysis of binary data is sometimes performed by fitting generalised linear models (GLMs). There were two reasons why GLMs were not applied to the data from this study. Firstly, the percentages used in the analysis were not true binary data, having been estimated from several other values; they could not be expressed as pairs of observed whole numbers from which the percentages had been calculated. Secondly, even advanced statistical software did not have the facility to incorporate sampling variance into such models.

The analysis produced mean percentages for each of the five times of day, averaged across the sampling dates. These means, however, were not simple arithmetic means, because percentages associated with larger
sampling variances were given less weight, and overall differences between sampling dates had to be accounted for in correcting for missing values.

The mean percentages were compared using a Wald statistic, which is analogous to the variance ratio used to compare treatments in ANOVA, though it has an asymptotic chi-squared distribution rather than an $F$ distribution. Comparisons between two particular times of day were made by calculating the least significant difference (LSD) at the 0.05 level for the associated pair of means. The LSDs varied slightly between different pairs of means because of the lack of balance in the data.

When analysing percentages or proportions, the usual assumption of homogeneity of variance is sometimes better met by applying the arcsine square root (also called angular) transformation to the data. However, none of the plots of residuals against fitted values demonstrated tangible improvement due to the transformation, and so it was not applied.

All calculations involved in estimating proportions and their associated variances, and all subsequent statistical analyses, were performed using the statistical package GenStat for Windows, 5th Edition (Lawes Agricultural Trust, IACR-Rothamsted).

### Example and discussion

The following example from our study illustrates the concepts and methods described above. Table 1 shows the results for pink-eared duck from the mid-afternoon sample on 19/5/1999. Graphical summaries of the results for all

<table>
<thead>
<tr>
<th>Section</th>
<th>Sub-sample</th>
<th>Feeding</th>
<th>Resting</th>
<th>Locomotion</th>
<th>Comfort movement</th>
<th>Other</th>
<th>Activity not observed</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>–</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>B1</td>
<td>1</td>
<td>1</td>
<td>72</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>B1</td>
<td>2</td>
<td>3</td>
<td>56</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>64</td>
</tr>
<tr>
<td>B1</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>763</td>
<td>763</td>
</tr>
<tr>
<td>B2</td>
<td>–</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>C</td>
<td>–</td>
<td>12</td>
<td>0</td>
<td>6</td>
<td>68</td>
<td>0</td>
<td>0</td>
<td>86</td>
</tr>
<tr>
<td>D</td>
<td>–</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>
species at all times of day (pooled across sampling dates) are presented in Hamilton et al. (in press).

The preliminary scan revealed that sections A, B2, C and D each had less than 100 individuals, and so in these sections all pink-eared ducks were observed and assigned to an activity class. Section B1 clearly had more than 100 individuals, and so was sub-sampled, as described previously. All pink-eared ducks observed in the two randomly selected sectors of B1 were assigned to an activity class. The total number of individuals in the section was estimated to be 900, arrived at by counting nine groups of about 100. The number of birds whose activity was not observed is found by subtraction: 900 – 73 – 64 = 763.

We now describe the calculations required to estimate the proportion of individuals in the pond undertaking a particular activity at this time, together with the associated sampling variance estimate. For the purpose of illustration, we will only consider ‘resting’, but the procedure applies in the same way to any activity. Steps 1, 2 and 3 below correspond to the same numbered steps given previously in ‘Estimation of proportions’.

1. In section B1, the number of individuals resting is summed across the two sub-samples, giving 72 + 56 = 128. The number not resting is summed similarly, giving 1 + 3 + 5 = 9. An estimate of the proportion resting is 128/(128+9) = 0.934, or 93.4%. An estimate of the number resting in the entire section is 0.934 × 900 = 840.9.

2. The variance associated with the estimated proportion, 0.934, is found from equation 1. Here, two ‘clusters’ of 73 and 64 individuals were selected from a population of 900. The values of the variables in the equation are \( N = 900; n = 73 + 64 = 137; a_1 = 72; a_2 = 56; m_1 = 73; m_2 = 64; k = 2; p = 0.934 \). Substituting these quantities into equation 1 gives \( v_r(p) = 0.00261 \). This variance is equivalent to a standard error of \( \sqrt{0.00261} = 0.051 \), or 5.1%. The estimated variance of the number of birds resting in section B1 is 0.00261 × 900^2 = 2110. The square root of this figure, 900 × 0.051 = 45.9, is the standard error associated with the estimated number resting in this section, 840.9. If we had assumed that a simple random sample was taken from the population, the estimated variance would have been given by the formula

\[
v_r(p) = \frac{N - n}{N} \frac{p(1 - p)}{n - 1}
\]

(3)
(Cochran, 1977, p. 52). For this example, \( v_r(p) = 0.000787 \), which is only about 30% of \( v_c(p) \). Such a gross underestimation of the variance occurs because the individuals within a sector were more homogeneous than would occur randomly. This ‘overdispersion’ is evident from the numbers in Table 1; for example, the proportion of birds resting tended to be either very high or low.

3. An estimate of the proportion of individuals resting across the entire pond is found from equation 2. In sections A, B2, C and D, \( N_j p_j \) simplifies to the number of birds observed to be resting; for example, in section A, \( N_j = 6 \) and \( p_j = 1/6 \), giving \( N_j p_j = 1 \). In section B1, \( N_j p_j \) is the figure calculated in step 1, \( 900 \times 0.934 = 840.9 \). So the estimated proportion for the pond is \( (1 + 840.9 + 0 + 0 + 1)/(6 + 900 + 3 + 86 + 3) = 842.9/998 \) = 0.845, or 84.5%.

The sampling variance associated with this proportion is equal to the variance in section B1 only, because the other sections had no sub-sampling. In B1, the estimated variance of the number of birds resting is 2110, so this figure applies also to the number resting in the entire pond (estimated to be 842.9). The estimated variance of the proportion of birds resting is \( 2110/998^2 = 0.00212 \), which is equivalent to a standard error of \( \sqrt{0.00212} = 0.046 \), or 4.6%.

These calculations therefore produced an estimate (84.5%) and an associated sampling variance \( (0.00212 \times 100^2 = 21.2) \) for the percentage of pink-eared ducks resting at mid-afternoon on 19/5/1999. The same steps were applied to the results from all other sampling occasions when pink-eared ducks were observed, resulting in a data set from which the effect of time of day on percentage of pink-eared ducks resting could be examined. The calculations were performed for all species \( \times \) activity combinations, but we will again focus on just this one for the purpose of illustration.

If the sampling variance is ignored, the percentages can be analysed by a standard ANOVA, with the means for the five times of day being compared, and sampling date modelled as a blocking factor. The results of an ANOVA applied to the percentages of pink-eared ducks resting are shown in Table 2.

The LSD has a range of values here because it depends on the number of missing observations for the times of day being compared, a missing observation being the result of observing no pink-eared ducks on that occasion. For example, to compare sunrise (8 missing observations out of
Table 2. Effect of time of day on the mean percentage of pink-eared ducks resting on Pond Nine; comparison between means was performed using ANOVA

<table>
<thead>
<tr>
<th>Time of day</th>
<th>Sunrise</th>
<th>Mid-morning</th>
<th>Midday</th>
<th>Mid-afternoon</th>
<th>Sunset</th>
<th>LSD (p = 0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% resting</td>
<td>10.8</td>
<td>59.4</td>
<td>60.8</td>
<td>45.8</td>
<td>31.4</td>
<td>12.5-13.2</td>
</tr>
</tbody>
</table>

41) with sunset (7 missing), the LSD is 12.5, while for midday (12 missing) versus mid-afternoon (10 missing) the LSD is 13.2.

The results of this ANOVA would suffice as a initial summary of the data, with the main pattern being clear. However, the analysis ignores the presence of sampling variance for some observations. This will, at least to some extent, result in underestimating the variability used to perform comparisons between means, and can give misleading estimates of the means themselves.

Sub-sampling was employed on 63 occasions out of a possible 159. The estimated sampling variance ranged from 0.04 to 914, with a median of 21.3 (the equivalent standard errors being 0.2%, 30.2% and 4.6% respectively). The sampling variance for the remaining 96 occasions was zero, because there was no sub-sampling. The wide range in these variances means that some proportions were estimated much less precisely than others. A standard ANOVA assumes that all the data have been collected or estimated with the same precision, i.e. they have equal sampling variances. In some circumstances a ‘weighted ANOVA’ can be used, with weights allocated to observations in inverse proportion to their estimated variances. But this was not possible here, because the observations that did not require sub-sampling had a sampling variance of zero. The more general and robust REML algorithm was able to give the appropriate weighting to each observation, as well as account for the substantial number of missing values. For these reasons, REML was used to analyse all the activity data from this study (see Hamilton et al., in press). The results of a REML analysis of the percentages of pink-eared ducks resting are shown in Table 3.

The LSD here is about 4% larger than the figure produced by a standard ANOVA. Using the ANOVA to compare means would slightly underestimate the variability, and hence slightly overstate the statistical significance.

The means shown in Table 3 also differ from those produced by ANOVA, for reasons explained previously. The effect of using REML is to increase
Table 3. Effect of time of day on the mean percentage of pink-eared ducks resting on Pond Nine; comparison between means was performed using REML

<table>
<thead>
<tr>
<th>Time of day</th>
<th>Sunrise</th>
<th>Mid-morning</th>
<th>Midday</th>
<th>Mid-afternoon</th>
<th>Sunset</th>
<th>LSD (p = 0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% resting</td>
<td>10.3</td>
<td>58.3</td>
<td>63.6</td>
<td>44.8</td>
<td>30.3</td>
<td>12.9-13.7</td>
</tr>
</tbody>
</table>

The estimated percentage resting at midday by about 3 percentage points, and to decrease it for the other times by about 1 percentage point. These differences are not substantial in this context, and the overall pattern of when the birds are likely to rest is very similar. However, in other situations the differences may be important, especially when taken in combination with different estimates of variability; the statistical significance of comparisons between means could alter considerably, leading to different conclusions from the study.

There are two main reasons why the means from the two types of analysis do not differ substantially in this example. One is that the size of the sampling variance is not related in any obvious way to the size of the estimated percentages \( r = 0.10 \) for these data. If, for example, the sampling variance is large when the estimated percentage is large, then accounting for the sampling variance in a REML analysis puts less weight on larger values, thus making the estimated means smaller. But if the means are affected in a much more haphazard fashion, the effect can be in either direction, and can easily even out, as it has to a large degree here. The other reason is that the sampling variance is not large compared to the unexplained, or residual, variation in the estimated percentages after accounting for differences between times of day and sampling dates. The residual mean square from the ANOVA is 664, and the corresponding figure from REML is 677. Compared to either value, the median non-zero sampling variance, 21.4, is very small. Even the largest sampling variance, 914, is only of similar magnitude — including it in the analysis results in only a 2-3 fold increase in the total variance for that observation.

The pattern we have observed here for pink-eared ducks resting, i.e. the variability slightly underestimated by a standard ANOVA, the means altered, but overall trends similar, occurred for most species \( \times \) activity combinations. One species for which the variability was more seriously underestimated
(up to around 12%) was hardhead, the only diving species in the study. Overall, the extra complication involved in using REML for the analysis may seem to be of little benefit in this situation, with most of the conclusions largely unaltered. But in other circumstances, it may be important that REML be used, to avoid overstating significance, or producing means which are misleading. Overstating the significance may occur when the sample size is small, because sampling variance increases with decreased sample size. Means which are misleading are likely to arise when the sampling variance is correlated with the size of the estimates. This may occur, for example, for species where social interaction varies with flock size.

The methods described in this paper, while having particular application to waterfowl activity budget studies, could be applied to scan sampling in other fields of study, or to any discipline where the sampling characteristics are similar. Some of the statistical analysis requires advanced statistical software, but the calculations (estimation of proportions and sampling variances) do not. We used Genstat for all calculations and analysis, as it was convenient to do everything in the one package, but the calculations could easily be done using a spreadsheet or other calculating device.

References


Patterson, H.D. & Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. — Biometrika 58, p. 545-554.


Social constraints on female mate preferences in mallards, *Anas platyrhynchos*, decrease offspring viability and mother productivity

CYNTHIA K. BLUHM* & PATRICIA ADAIR GOWATY†

*Delta Waterfowl and Wetlands Research Station
†Institute of Ecology, University of Georgia

(Received 19 November 2002; initial acceptance 4 February 2003; final acceptance 28 January 2004; published online 12 September 2004; MS. number: A9491R)

Constraints theory says that when individuals reproduce with nonpreferred partners, they will have offspring of lower viability than when individuals reproduce with preferred partners. To test this, we performed mate preference trials free of experimenter biases about what traits in males make them attractive to females. We eliminated social interactions that could manipulate or obscure females' choices, including interactions between males, between females, and between the sexes. We then randomly placed each female with either her nonpreferred or preferred male in breeding trials to determine how components of fitness for breeders and offspring viability varied. Offspring viability (the percentage of eggs surviving to 45 days posthatch) and mother productivity (mean number of offspring at 45 days posthatch) were significantly lower when mothers reproduced with their nonpreferred partner than with their preferred partners. This study adds to the growing list of species in which reproduction with nonpreferred partners is known to be costly to breeder and offspring fitness. It is among the few studies of birds and the only study of a species with forced copulation showing fitness deficits for offspring and mothers from enforced reproduction with nonpreferred males.

2004 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd.

Constraints theory (Gowaty 1996, 1997, 2003; Gowaty & Buschhaus 1998) says that variation among females in their abilities to minimize or avoid altogether the effects of ecological or social limitations on their reproductive decisions results in female–female sexual selection. Forced copulation is a dramatic example of such a constraint. However, other, less dramatic-to-observe constraints with strong effects are common. If variation among individuals in their abilities to minimize or avoid constraints on their reproductive decisions is also common, selection due to constraints will be common. For example, dispersal limitation may constrain some females to reproduce with kin that are nonpreferred mates (Ryan & Altmann 2001), while other females, more energetic or vigorous perhaps, may escape local limits and mate with out-bred (Penn & Potts 1999) preferred partners. Predation risk causes previously choosy females to mate indiscriminately (Hedrick & Dill 1993; Dill et al. 1999; Hedrick 2000; Johnson & Basolo 2003). If some females are able to avoid mating under predation risk, but others are not or do so less reliably, female–female reproductive competition over the quality of mates and offspring may commonly result. Pre- and postcopulatory mate guarding may keep other males away from females, but mate guarding just as reliably limits the opportunities of females to seek alternate mates (Gowaty 1985; Gowaty et al. 1989; Gowaty & Bridges 1991). If there is variation among females in their abilities to minimize lost opportunities to express their preferences or mate polyandrously due to pre- or postcopulatory guarding, female–female reproductive competition over the quality of mates and offspring may result (Gowaty 1997; Foerster et al. 2003). As preferences for traits in heterospecifics have demonstrated (Basolo 1990; Ryan et al. 1990), even attractive male traits may manipulate females into reproductive decisions deleterious to their fitness or the fitness of their offspring. If there is variation
among females in their abilities to escape exploitation via sensory biasing, female–female reproductive competition over the quality of mates and offspring may result.

Constraints theory assumes that both sexes have mate preferences that are shaped by variation in offspring viability. However, constraints theory is silent about the cues that choosers use, although it is consistent with use of multiple sensory modalities for mate choice.

Here we report the results of tests, using mallards, of the prediction that offspring whose mothers reproduce with males they do not prefer (NP) have significantly lower viability than offspring whose mothers reproduce with males they do prefer (P). This study was a companion to other studies designed to test simultaneously the prediction that offspring viability suffers when breeders reproduce with NP mates. Besides the ducks that we report on here, the study species included two species of insect (Moore et al. 2001, 2003; W. W. Anderson, P. A. Gowaty & Y.-K. Kim, unpublished data), two fish species (Sandvik et al. 2000; J. F. Downhower & M. L. Matsui, unpublished data), and a mammal species (Drickamer et al. 2000, 2003; Gowaty et al. 2003). Constraints theory in the context of life histories (Gowaty 1997) suggests that types of impediments (temporal, spatial, and social constraints) on reproductive decision making probably vary in strength and significance depending on taxon, ecological circumstances and social systems. Despite the expected variation in constraints that is likely to be experienced by species in our companion studies, we assumed that in nature all of these species experience some constraints that probably affect reproductive options of some individuals and subsequent decisions of breeders. Thus, our studies shared a common design that tested the fitness effects on breeders and offspring when breeders were either experimentally constrained to breed with partners they did not prefer or allowed to reproduce with partners they did prefer. Note that our question makes an important distinction between nonpreferred and random partners. Thus, our experimental design differs from those in which the fitness of choosers that reproduce with partners they prefer is compared to the fitness of choosers that reproduce with random partners. This is partly because our design was not motivated by a desire to explain the evolution of traits in males or sire effects on offspring (Reynolds & Gross 1992; Petrie 1994), but by a desire to explain the effects of constraints on breeding females. We examined the costs to females of mating with nonpreferred partners relative to preferred partners.

In nature, female mallards may often be limited to reproduction with males they do not prefer. First, males guard females’ movements during egg laying (Cheng et al. 1982; Titman 1983), perhaps to decrease the probability that females will seek extrapair copulations. Second, male aggression against females may condition or manipulate (constrain) females’ subsequent behaviour for the benefit of males (Gowaty & Buschhaus 1998). Third, some males coerce copulation with some extrapair females with aggressive force (McKinney et al. 1983) during the winter prebreeding period (McKinney et al. 1983) and also during egg laying (Cheng et al. 1982; Titman 1983). Some forced copulations result in successful fertilization (Evarts & Williams 1987). Are there fitness consequences for females and their offspring from coerced or manipulated reproduction?

We took a components-of-fitness approach to answer this question, because constraints theory makes predictions about offspring viability. Offspring viability is a component of offspring fitness. But, offspring viability also may affect the fitness of mothers: lower offspring viability may result in lower productivity (number of offspring that reach the age of reproduction) of mothers that reproduce with NP compared to P males. Offspring viability also may have other effects on mothers’ fitness. Mothers reproducing with NP males may adjust their fecundity (increase the number of eggs they lay) in an attempt to compensate for likely offspring viability deficits (Gowaty 2003; Bluhm & Gowaty 2004). Mothers mated with NP males may have lower fecundity because of zygote loss before eggs are laid. Thus, for species in which each individual has two parents, it is usually impossible to completely partition components of fitness separately for mothers, fathers and offspring. What is unambiguous, however, is that whenever offspring viability varies between mothers, it is a selection pressure acting on their behaviour and physiology.

We estimated offspring viability for successful mothers (see below) in terms of offspring survival from the egg stage to independence. We measured the productivity of mothers as the number of independent offspring per female, and the fecundity of mothers as the number of eggs laid per female. In this and in our companion studies, we also measured other fitness components of mothers including fertility (the number of bouts of reproduction). We used offspring viability as an estimate of an offspring’s viability throughout life.

Because there may be many reasons for offspring inviability, including sperm insufficiencies, variation in the deposition of egg proteins around fertilized eggs, or maternal physiological or behavioural pathology, we were most interested in the variation in offspring viability of mothers that had at least one offspring fledge. We also controlled for variation in mothers’ behaviour by incubating eggs in an incubator and raising precocial young in brood pens. Thus, significantly higher offspring viability of successful mothers (those with at least one fledged offspring; reviewed in Drilling et al. 2002) that reproduce with P versus NP males would constitute strong support for the constraints hypothesis. Lower offspring viability among successful mothers is also likely to result in fewer offspring surviving to reproductive age (productivity), and thus, the constraints hypothesis predicts that, for successful mothers, reproduction with NP partners has negative effects on their fitness.

**METHODS**

**Study Species**

Details of mallard natural history and reproduction are in Drilling et al. (2002). Seasonal pair bonds form on the wintering grounds, and all females arrive paired on the
breeding grounds. Females lay large clutches of 8–12 eggs. Young are precocial and capable of flight at 6 weeks.

We used 222 mallard adults hatched from eggs collected from nests of wild mallards in southern Manitoba. We held them at Delta Waterfowl and Wetlands Research Station in Manitoba in mixed-sex groups until fledging age at 6 weeks old, when we placed them in same-sex flocks within sight and sound of mixed-sex and opposite-sex flocks. We housed flocks outdoors from August 1997 to mid-October 1997 and indoors from mid-October 1997 to April 1998 in a modern duck-holding facility with natural photoperiod. We made regular preliminary observations of same-sex groups before randomly drawing individuals for use in arena preference tests (see below). In many captive groups, male mallards sometimes form courting groups (C.K.B., personal observation). No males showed evidence of same-sex affiliative display or courting, so we included all males in arena tests.

Arena Preference Tests

We designed arena preference tests specifically to be free of (1) experimenter biases about what phenotypic characteristics of males make them attractive to females and (2) typical social interactions such as within-sex contests and between-sex coercion (Partridge 1980, 1981; Kingett et al. 1981; McKinney 1992). As in our companion studies (cited earlier), we picked males for inclusion in these trials at random with respect to phenotypic variation obvious to us. Thus, if male traits are sensory traps that can manipulate females, our inclusion of males at random with respect to their phenotypic variation probably randomized these effects. Although our use of arenas prevented within-sex contests and between-sex coercion, it allowed females to choose between males using all other sensory channels, and it allowed us to infer female preferences based on an ecologically reliable indicator: the amount of time that a female spent near each male.

In the mallard tests, we used three males. We usually used the same set of three stimulus males in three consecutive preference trials with different females. This allowed us to test whether individual males were consistently NP or P.

All experimental females were virgins, either in their first or second year. We used two age classes of virgins because not enough of one age class was available for the entire experiment, and because first-year virgins do not always breed (C.K.B., personal observation). We conducted 105 preference trials from 11 March to 24 April 1998. In 87 trials, females met our criteria (see below) for showing a preference. Twelve females were in two trials; 81 in one trial only. Brothers were never in preference tests with their sisters. All the males (N = 121) were second-year birds.

The indoor preference arena was a screened-in, four-chambered, rectangular pool (water depth 12–15 cm) that measured 1.84 × 4.22 × 3 m high and was under natural photoperiod. Three of the chambers occupied one-fifth of the arena and were visually isolated from each other using opaque wood for sides. One chamber occupied the other four-fifths of the arena and allowed visual access to all three smaller chambers. To begin a trial, we placed three males into the smaller chambers, then placed a focal female in the large chamber. Preference trials lasted for 1 h and were conducted during 0730–1100 hours and 1600–1830 hours. We scored a female as having made a preference if she spent at least 50% (30 min) of the total test time within 1 m of any male and 25 min within 1 m of one male. We defined the male with whom she spent at least 25 min as the preferred (P) male. All females who met these criteria spent at least some time with the remaining two males, so we assigned the one with whom she spent the least time the NP male.

Breeding Tests

To eliminate residual effects of previous reproduction with one partner on later reproduction with others, we used a between-female design here and in simultaneous tests (cited above) with other internally fertilizing species. We placed each pair of birds into outdoor breeding compartments on the same day, 26 April, to control for the time available for breeding for each pair.

For females showing a preference, we used a stratified, random rule for assigning them to separate compartments containing their NP (N = 31) or P (N = 41) male. No male was with more than one female. Because only 72 breeding pens were available, we used only 72 of the 87 females that met our behavioural criteria for showing a preference. Within each age group (first- and second-year virgins), we ranked each female by the time she spent in front of her most preferred male; then, we made alternating assignments of females to breeding compartments with their NP or P male. This assignment rule controlled for the distribution of ‘preference strength’ between the two experimental breeding categories. For second-year females (N = 49), 21 were with their NP male and 28 (57%) with their P male. For first-year females (N = 23), 10 were with their NP male and 13 (56.5%) with their P male. We weighed females on the day we placed them in the breeding compartments and provided all animals with ad libitum food and water throughout the study. Females remained with their assigned male for first and subsequent breeding attempts (i.e. throughout the breeding season).

Adult females weighed significantly more (1078.5 ± 11.5 g) than yearling females (977.8 ± 11 g; t test: t35 = 5.48, P < 0.0001). However, there were no differences in the weight of females assigned to breeding compartments with NP (1044 ± 15.2 g) or P (1046.9 ± 18 g; t35 = −0.094, P > 0.926) males.

Components of Fitness

We refer to individuals that hatched as ‘ducklings’, and to those that were at least 42 days old and capable of flight as ‘fledglings’.

A technician, who was unaware of the experimental treatment of breeders, made all observations on eggs and ducklings. Each morning, he checked nestboxes, numbered eggs in sequence, and returned them to the nest. We identified females that failed to lay any eggs as nonlayers;
for layers, we noted the total number of eggs laid. We incubated eggs in a forced-air incubator. After hatching, we web-tagged each duckling (number 5 fingerling tags, National Band and Tag Co., Newport, Kentucky, U.S.A.), and reared mixed-sex broods separately in indoor enclosures. Three times a day we censused enclosures and removed any dead ducklings. There was no negative effect of creche size on offspring survival, the correlation between initial creche size and the number of offspring that survived was positive. After 30 days posthatch we moved the ducklings into much larger mixed-brood flight pens. We released each fledgling to the wild at the Delta Station in Manitoba, when they were 50 days old or older.

**Statistical Analysis**

We operationally defined mallard productivity as the number of offspring that survived to 50 days of age, which is when fledglings are often independent in the wild (Bluhm 1985; Afton & Paulus 1992; Drilling et al. 2002). We report variation in the number of clutches laid (fertility), number of eggs laid (fecundity), number of eggs that hatched (the number of ducklings at hatch), the number of offspring that survived to the release date (productivity), as well as offspring viability ((productivity/fecundity) × 100) from laying to hatching, from laying to 50 days posthatching and from hatching to 50 days posthatching.

Initially, we analysed the effects of mother age (first year or second year) and weight on components of offspring fitness. Because we found no effect of age or weight, and no interaction between these variables on fertility (number of clutches), fecundity, productivity, or offspring viability over any interval, we performed t tests and chi-square tests, using only mating type (with NP or P male) as the main effect, in all subsequent analyses. We analysed fitness components of mothers (fecundity and productivity) and offspring viability over various intervals. Because not all mothers produced eggs or ducklings, we report mothers’ fitness components in several ways: based on all females (N = 72), only females that laid eggs (N = 49), only females that had at least one egg hatch (N = 40) and only females that had at least one flying offspring (N = 40; i.e. successful). The most conservative measures of the offspring viability prediction from the constraints hypothesis are those that control for possible sperm insufficiency and maternal pathology. Therefore, we consider the most reliable measure of the predicted variation in offspring viability to be those mothers having at least one duckling reach independence.

We performed analyses on arcsine square-root transformed variables that did not meet the assumption of normality. We set significance probability at 0.05.

**RESULTS**

**Offspring Viability and Mother Productivity**

Offspring of successful mothers (those with one independent offspring) that reproduced with NP partners had significantly lower egg-to-independence viability compared with offspring whose mothers reproduced with P partners (Table 1). Number of independent offspring and number of independent offspring/clutch (mother productivity) were also significantly lower when mothers reproduced with their NP rather than their P partners (Table 1).

**Fertility and Fecundity of Mothers**

There were no differences in the proportion of females by mating type that laid eggs (NP = 23/32; P = 26/41), produced one duckling (NP = 20/23; P = 21/26), or one fledgling (NP = 20/20; P = 21/21). Fertility (the number of clutches produced/female) and fecundity (expressed as the total number of eggs laid and as the total number laid per clutch; Table 2) were not significantly different for females that reproduced with their NP or P partners. Nor were there any significant differences by mating type in the number of eggs that hatched or the percentage of fertile eggs (Table 2).

**No Best Males**

The time that mothers spent with their assigned males during preference trials correlated significantly with offspring viability from laying to 50 days posthatching (Pearson correlation: r = 0.33, P = 0.038). Yet, overall, there were no ‘best or consensus males’. Sixty-three males that were most preferred in one test included 38 that were least preferred in another. Likewise, 60 males that were least preferred in one preference test were most preferred in 35 others. Other analyses also provided evidence against the hypothesis that there is a ‘single best male that all females prefer’. First, there was no association between the number of times a male was in a choice test and his preference ranking. In this analysis there were three preference ranks and six categories of males based on the number of times that a given male was used in preference tests (1–5 and >5; chi-square: χ² = 5.8, NS). Second, a maximum likelihood analysis of the trinomial probabilities that a male was most preferred, least preferred, or middling preferred showed that over all trials, probabilities were almost exactly evenly one-third, which would not be expected under a model in which a single ‘best male’ is preferred by all or most females.

**DISCUSSION**

Offspring of mothers that reproduced with NP males had lower viability and decreased survivorship from laying to 50 days posthatching compared with offspring of mothers that reproduced with P males. Our study controlled for variation in maternal behaviour towards eggs and ducklings, as well as for maternal or paternal pathologies that might have contributed to mothers’ failures to lay fertile eggs. In addition, successful mothers that reproduced with NP males had significantly fewer independent offspring (at the age of release) than mothers that reproduced with P males. These results are similar to studies of Mus musculus (Drickamer et al. 2000), Drosophila pseudoobscura (Gowaty
et al. 2002; Anderson et al., unpublished data), *Peromyscus polionotus rhoadsi* (Ryan & Altmann 2001) and *Oryzias latipes* (J. F. Downhower & M. L. Matsui, unpublished data), all of which shared important design features with our study. Our results are also similar to those of other studies in which the design features confound female and male choice but show that pairs with some choice have higher reproductive success than randomly constituted pairs (McLain 1998). Our results are also similar to those in studies designed to examine sire effects of relatively attractive males (Petrie 1994), but which did not take into account individual female preferences for individual males.

Between-female variation in productivity due to variation in females’ abilities to avoid social constraints is likely to be a strong selection pressure on maternal behaviour and physiology. Thus, we predicted that mothers that reproduce with males they do not prefer will attempt to compensate (Gowaty 2003) for deficits in offspring viability. Elsewhere (Bluhm & Gowaty 2004) we report data consistent with compensation for offspring viability deficits in mallards.

There was a positive correlation between the time a female spent with assigned males during preference tests and egg-to-fledgling viability. These results are similar to experimental results on *Drosophila pseudoobscura*, which showed that approaches to opposite-sex partners in dyadic encounters are significantly associated with breeder productivity (Gowaty et al. 2002). They are also similar to data on stink bugs *Nezara viridula* (McLain 1998) in which premating duration of choosing bugs is significantly negatively correlated with reproductive success. We are unaware of similar data on other vertebrate species.

In contrast to previous avian mate preferences studies, we assigned males to preference tests at random with respect to phenotypic differences, yet females still preferred some males over others. Because we did not measure male trait variation, we cannot comment on the cue basis of female mate preferences. It seems unlikely, however, that females use plumage variation or vigour indicators alone, because in preference tests using the same males, different females seldom preferred the same male. Furthermore, the short interval over which we performed preference trials was unlikely to have been

---

### Table 1. Fitness components of successful mothers (those with at least one independent offspring) and their offspring

<table>
<thead>
<tr>
<th>Fitness component</th>
<th>Mating type</th>
<th>NP (N)</th>
<th>P (N)</th>
<th>t</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of eggs laid</td>
<td></td>
<td>19.8±2.24 (20)</td>
<td>17.9±1.6 (21)</td>
<td>0.585</td>
<td>39</td>
<td>0.563</td>
</tr>
<tr>
<td>Number of ducklings (at hatch)</td>
<td></td>
<td>9.3±1.2 (20)</td>
<td>13±1.2 (21)</td>
<td>−2.188</td>
<td>39</td>
<td>0.034</td>
</tr>
<tr>
<td>Number of fledglings at 50 days/ clutch</td>
<td></td>
<td>5.5±0.61 (20)</td>
<td>7.2±0.59 (21)</td>
<td>−2.069</td>
<td>39</td>
<td>0.0452</td>
</tr>
<tr>
<td>Number of fledglings surviving to 50 days</td>
<td></td>
<td>8.5±1.2 (20)</td>
<td>12.3±1.14 (21)</td>
<td>−2.374</td>
<td>39</td>
<td>0.0025</td>
</tr>
<tr>
<td>% Ducklings (from hatching to 50 days viability)</td>
<td></td>
<td>91.3±2.4 (20)</td>
<td>96±2.4 (21)</td>
<td>−1.403</td>
<td>39</td>
<td>0.17</td>
</tr>
<tr>
<td>% Eggs surviving to 50 days (from laying to 50 days viability)</td>
<td></td>
<td>51±5.4 (20)</td>
<td>69±5.2 (21)</td>
<td>−2.382</td>
<td>39</td>
<td>0.022</td>
</tr>
</tbody>
</table>

---

### Table 2. Fitness components (means ± SE) for females paired with their nonpreferred (NP) or preferred (P) males. Fitness components were calculated using means/mother

<table>
<thead>
<tr>
<th>Fitness component</th>
<th>NP (N)</th>
<th>P (N)</th>
<th>t</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of clutches/female</td>
<td>1.3±0.18 (31)</td>
<td>1.02±0.15 (41)</td>
<td>1.14</td>
<td>70</td>
<td>0.26</td>
</tr>
<tr>
<td>Number of eggs/female</td>
<td>13.9±1.9 (31)</td>
<td>10.7±1.7 (41)</td>
<td>−0.947</td>
<td>70</td>
<td>0.35</td>
</tr>
<tr>
<td>Clutches/layers female</td>
<td>1.74±0.15 (23)</td>
<td>1.62±0.14 (26)</td>
<td>0.598</td>
<td>47</td>
<td>0.56</td>
</tr>
<tr>
<td>Number of eggs/layers female</td>
<td>18.8±1.8 (23)</td>
<td>16.9±1.6 (26)</td>
<td>0.789</td>
<td>47</td>
<td>0.44</td>
</tr>
<tr>
<td>Number of eggs laid per clutch</td>
<td>10.6±0.40 (23)</td>
<td>10.7±0.36 (26)</td>
<td>−0.106</td>
<td>47</td>
<td>0.92</td>
</tr>
<tr>
<td>% Fertile eggs/layers female</td>
<td>68±7.1 (23)</td>
<td>69.1±6.7 (26)</td>
<td>−0.248</td>
<td>47</td>
<td>0.81</td>
</tr>
<tr>
<td>Number of ducklings per clutch per layer female</td>
<td>5.13±0.74 (23)</td>
<td>6.1±0.7 (26)</td>
<td>−0.958</td>
<td>47</td>
<td>0.395</td>
</tr>
<tr>
<td>(i.e. number of eggs hatched)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of ducklings/layers female</td>
<td>8.1±1.4 (23)</td>
<td>10.5±1.3 (26)</td>
<td>−1.276</td>
<td>47</td>
<td>0.21</td>
</tr>
<tr>
<td>(i.e. number of eggs hatched)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Egg to hatching viability/layer female (%)</td>
<td>47.6±6.8 (23)</td>
<td>57.9±6.4 (26)</td>
<td>−1.106</td>
<td>47</td>
<td>0.275</td>
</tr>
</tbody>
</table>
long enough for male vigour or plumage traits to have changed.

Offspring viability benefits from mating with P males could come from specific alleles from their fathers that increase resistance to particular diseases, as the ‘good genes’ hypotheses argue (Hamilton & Zuk 1982; Brown 1997; Wedekind 1999). If parasite and pathogen populations are evolving rapidly, as many have argued (Van Valen 1973; Dybdahl & Lively 1998; Lively & Dybdahl 2000), it is likely that alleles providing defence against parental generation pathogens often will not work against offspring generation pathogens (Hamilton & Zuk 1982). If this is so, offspring viability will most likely be a function of alleles from both parents, so that whenever females in populations vary, their mate preferences are likely to be self-referential, meaning that who is an optimal mate for one female is unlikely to be optimal for all others (Gowaty 2003). This suggests that it is improbable that there are a few best males or a single best male that all females prefer. In all but inbred populations, it predicts that specific male(s) preferred by a given female will be non-preferred by others. Our data are consistent with hypotheses that take into account variation in both parents simultaneously. Whether male phenotypic traits affect individual female preferences in mallards, and whether these traits evolve because they signal a male’s possession of heritable disease-fighting alleles or complementarity with females’ disease fighting alleles remain open questions. It would also be interesting to know whether MHC complementarity with males affects female preference in mallards.

The modified differential allocation hypothesis (MDAH) (Sheldon 2000) says that viability benefits for offspring of P males could result if females bias investment in them. Elsewhere (Bluhm & Gowaty 2004) we report variation in egg and duckling mass and growth to test, for the first time, the contrasting predictions of the MDAH and the compensation hypothesis (Gowaty 2003). We found (1) a positive relationship between the number of eggs and the mean size of eggs per mother, indicating that there is no trade-off between the number and quality of eggs; (2) there was no support for the MDAH, but there was for the compensation hypothesis; (3) significant heterogeneity among females by age showed that second-year mothers laid heavier eggs when reproducing with NP than P males, but first-year mothers, whose egg sizes did not vary significantly with the preference status of their mates, did not; (4) by the time of independence, there were no differences in size and condition of offspring of second-year mothers, but there were for first-year mothers. At independence, the offspring of first-year mothers that reproduced with NP males weighed significantly less and were in poorer condition than the offspring of first-year mothers that reproduced with their P males.

We are currently unable to comment on the possibility that display rates and the frequency of forced copulations vary when females are with NP and P partners, because modest financial support limited us to essential observations of fitness components and offspring characteristics. Nor can we comment on variation in mothers’ incubation behaviour or other effects of female and male interactions with eggs or ducklings, because our design eliminated these potential parental effects. Such observations should be a part of future studies of fitness components and mate preferences in mallards.

Our data are the first to show that reproduction with nonpreferred partners in a species with forced copulation is costly to the fitness of offspring and mothers. Lower productivity from NP pairings should favour female resistance to copulation and fertilization. If fertilization is successful, lower productivity should favour attempted compensation (Gowaty 2003), which might be behavioural or physiological (Bluhm & Gowaty 2004). The fitness differences between females that reproduced with their NP and P males suggest why females resist copulation with some males (i.e. why some copulations appear to be ‘forced’, Gowaty 1997). Future behavioural data could differentiate the hypotheses that forced copulation results because males are NP versus the hypothesis that forced copulation results because females use resistance to assess male quality (Hoi 1997). If female resistance is as common or more common when females are paired with P than NP males, the hypothesis that females use resistance to copulations to assess male quality would be strengthened. If forced copulations are by males that females do not prefer, then forced copulations should be less successful than nonforced copulations in achieving immediate fertilization success (Gowaty & Buschhaus 1998), as shown in zebra finches, Taeniopygia guttata (Burley et al. 1996).

Acknowledgments

We thank Shawn Tully for expert technical assistance in collecting egg and duckling data. We offer special thanks to Kevin Ward and Stephen P. Hubbell for help and advice, W. W. Anderson, L. C. Drickamer, Steve Hubbell, Allen Moore and our students for helpful comments, and Randy Nelson for his encouragement. A grant from the NSF supported this work. The University of Georgia Animal Care Committee, the Animal Care Committee of Delta Waterfowl and Wetlands Research Station, approved experimental protocols. We handled all birds under migratory bird license (WS-M51, CWS98-M005, CWS99-M001) to C.K.B. from the Canadian Wildlife Service.

References


Reproductive compensation for offspring viability deficits by female mallards, *Anas platyrhynchos*

CYNTHIA K. BLUHM* & PATRICIA ADAIR GOWATY†

*Delta Waterfowl and Wetlands Research Station
†Institute of Ecology, University of Georgia

(Received 19 November 2002; initial acceptance 20 February 2003; final acceptance 28 January 2004; published online 12 September 2004; MS. number: A9492R)

The compensation hypothesis (CH) says that when impediments to the expression of mate preferences occur so that reproduction is with nonpreferred (NP) partners, mothers with NP mates compensate for lower-viability offspring through enhanced investment. Previously, we tested an assumption of the CH and showed that offspring viability was significantly lower when mothers reproduced with NP mates. Here we report tests of the CH. Egg mass differed significantly by mother agemate category. Second-year mothers in NP pairings laid significantly heavier eggs than second-year mothers in preferred (P) pairings; mean egg mass of first-year mothers with NP or P mates did not differ. However, mean egg mass per mother and mean number of eggs per mother were positively and significantly related. Thus, there was no trade-off between egg size and egg number for any category of mother. By fledging age, duckling performance, quality and condition were significantly lower for first-year mothers with NP mates compared with those for mothers in the other categories. Second-year, but not first-year mothers, successfully compensated for predictable deficits in offspring viability from reproduction with NP males. This is the first study showing that mothers compensate for viability deficits in offspring resulting from reproduction with NP partners.

2004 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd.

What should individuals do when in enforced monogamy with partners they do not prefer? If offspring viability is predictably lower with nonpreferred (NP) compared with preferred (P) partners, the compensation hypothesis says selection will favour individuals with NP mates compensating or attempting to compensate for offspring viability deficits through enhanced investment (Gowaty 2003). Lower breeder productivity (the number of offspring surviving to independence) will favour compensation because of lower offspring viability (Gowaty 2003). The compensation hypothesis predicts that females reproducing with NP mates should increase extrapair mating behaviour, fertility, fecundity and/or maternal investments in ways that are likely to enhance offspring health.

Mallard females reproducing with NP mates produce offspring of significantly lower viability (percentage of eggs surviving to independence) than females reproducing with P mates (Bluhm & Gowaty 2004). Therefore, mallards fit the assumption of the CH, and here we test predictions of the CH.

If males that are ‘attractive’ in groups are also those preferred by individual females, our data also address the modified differential allocation hypothesis (MDAH). Burley’s (1988) differential allocation hypothesis, on which the MDAH is based, posits that, in species with biparental care, the relative attractiveness of partners affects levels of relative parental investment and progeny sex ratios. More paternal than maternal provisioning should occur when a mother is relatively more attractive than a father and vice versa, and sex allocation should be biased towards the sex of the more attractive parent. Burley’s strict hypothesis does not apply to mallards, because parental care is by mothers only. However, others (Sheldon 2000) modified Burley’s hypothesis to predict that whenever females reproduce with attractive males, they increase their investment in offspring, by laying larger eggs, and, in some cases overproducing sons. Assuming larger eggs indicate greater maternal investments, the MDAH predicts that females that reproduce
with their P mates will lay heavier eggs and produce offspring of superior performance, quality or condition than females that reproduce with NP mates. The CH predicts that females that mate with NP males will lay heavier eggs than females that mate with P males and, if compensation from NP pairings is unsuccessful, then at independence, offspring performance, quality or condition will be less than that from P pairings. If compensation is successful, however, performance, quality or condition of offspring from NP and P mates will not differ (Gowaty 2003). If no differences exist in the quality of eggs laid by females by mate type, neither the MDAH nor the CH would be supported, even if fledgling performance, quality or condition varies.

**METHODS**

**Study Species and Study Sites**

Mallard natural history and breeding biology are reviewed in Drilling et al. (2002). We used breeders reared from eggs collected in the wild (Bluhm & Gowaty 2004) from farmers’ fields in Manitoba. All experiments were conducted at the Delta Waterfowl and Wetlands Research Station in Manitoba.

**Breeding Trials with NP and P Mates**

To test individual female preferences, we conducted arena tests and measured relative time in proximity of test females to different males, each in a different cell of an indoor pond. The arena design allowed females to see, hear and smell the males simultaneously, but the males could not see or interact with each other, nor could they touch the females. A comprehensive description of preference tests and behavioural criteria for showing a preference are in Bluhm & Gowaty (2004). To advance to breeding trials a female had to indicate a preference by associating with one male more than the others, because our interest was in fitness variation when females were limited to reproduction with partners known to be less preferred by them than P males. We compared fitness outcomes for females in enforced monogamy with either their most NP or P male.

After preference testing, we used a stratified random rule, and placed 72 first- and second-year virgin females alternately with their NP or P male in breeding trials on the same day in late April, controlling social exposure to mates, the time available for breeding, and the distribution of subjects by age and mass evenly to treatments (Bluhm & Gowaty 2004). Twenty-one second-year virgins (N = 49) were with their NP male and 28 with their P male. Ten first-year virgins (N = 23) were with their NP male and 13 with their P male. Forty-nine females produced 82 clutches and laid 870 eggs, of which 459 hatched and 430 reached independence (≥50 days posthatch).

**Egg and Duckling Masses**

A technician unaware of the experimental treatment of subjects made all observations on eggs and ducklings. He numbered new eggs in sequence on the morning that the hens laid them. After 5 days with no new eggs, he removed the eggs and replaced the nesting material. He weighed eggs with a Mettler digital pan balance and measured their length and width with vernier calipers. He washed all eggs in chlorinated washing solution to remove surface bacteria. Eggs developed in a forced draft incubator (Humidaire Incubator Co., New Madison, Ohio, U.S.A.). At 23 days, we put the eggs in individual compartments of a hatcher (Humidaire Incubator Co.). We cracked open all eggs that failed to develop after an additional 14 days, and coded them as fertile, infertile, or unknown. We uniquely web-tagged all ducklings at hatch (number 5 fingerling tags, National Band and Tag Co., Newport, Kentucky, U.S.A.). Until they were 30 days old, ducklings were in créches with their siblings in randomly assigned indoor enclosures that were visually isolated from other créches. However, ducklings in créches could hear and smell adults and ducklings in other créches. At 30 days we placed all ducklings from each créche into much larger, outdoor flight pens holding other créches. We weighed ducklings at hatch (N = 455) and at 15 days (N = 443), 30 days (N = 435) and 45 days (N = 433) posthatch. At 45 days posthatch, when ducklings normally fledge, we placed numbered aluminium leg bands on each duckling and measured their tarsi using vernier calipers. We released them to the wild at 47–50 days posthatch.

Throughout, ‘egg mass’ refers to fresh egg mass, calculated from Hoyt’s (1979) equation, \( W = K_w \times LB^2 \), where \( W \) is fresh weight, \( K_w \) is a species specific coefficient = 0.56 for mallards, \( L \) is length and \( B \) is breadth or maximum diameter. We estimated mean offspring condition per mother as the residual from a regression of mean mass of offspring per mother on mean tarsus length of offspring per mother.

**Statistical Analyses**

Because reproductive output of mallards is often significantly lower for first-year than second-year females, we included female age in all initial analyses. Mallard egg mass varies by clutch number (Batt & Prince 1978), so we also examined effects of clutch number (whether first, second, or third clutch laid by a given female) on mean egg mass and mass at hatching and when ducklings were 15, 30 and 45 days old. Only egg mass varied by clutch number. However, the mean number of clutches did not vary significantly by female age (first-year = 1.14 clutches; second-year = 1.17 clutches), mate type (NP = 1.27; P = 1.04), or their interaction (first-year with NP = 1.2 ± 0.3; first-year with P = 1.1 ± 0.3; second-year with NP = 1.3 ± 0.2; second-year with P = 1.0 ± 0.19).

Thus, in further analyses, we averaged all variables over clutches.

An ANOVA of egg mass with factors mother’s age (\( P = 0.1862 \)), mate type (\( P = 0.8795 \)), and their interaction
term \((P = 0.0287)\) showed that only the interaction term was significant. This result indicated a violation of the linearity assumption in ANOVA, so in further analyses of means per mother, we used single factor ANOVA with four categories: (1) first-year females with NP males (1-NP), (2) first-year females with P males (1-P), (3) second-year females with NP males (2-NP), and (4) second-year females with P males (2-P). For all significant one-way ANOVAs, we performed post hoc Student’s \(t\) tests to test for statistically significant differences between pairs of least square means. We transformed sex ratios using the arcsine square-root transformation to achieve normality before analysis of sex ratio variation. We report means ± SE and considered results significant when \(P \leq 0.05\).

We handled all mallards and eggs under migratory bird license (WS-MS1, CWS98-M005, CWS99-M001) to C.K.B. from the Canadian Wildlife Service according to protocols approved by the Animal Care Committees of the University of Georgia and the Delta Waterfowl and Wetlands Research Station.

RESULTS

Mass of Mothers and Their Eggs

Second-year mothers weighed significantly more than first-year mothers \((F_{3,41} = 8.8529, P < 0.0001)\), however, the mean mass of first- and second-year mothers in each mating type did not differ (Fig. 1a). Mother’s mass had no significant functional relationship with egg mass \((F = 0.0643, r^2 = 0.0014, \text{NS}; \text{Fig. 1b})\). Number of eggs per mother and mean egg mass per mother were significantly positively correlated; the residuals by female category were not significant \((F_{3,45} = 2.3965, \text{NS})\), and separate regression analyses by mother category were either nonsignificant \((1-P: r^2 = 0.0639, N = 8; 2-P: r^2 = 0.064739, N = 18; 1-NP: r^2 = 0.541968, N = 7)\) or significantly positive \((2-NP: r^2 = 0.411823, N = 16, P < 0.0075)\). The data, therefore, offer no support for trade-offs between egg number and size (Fig. 1c): the more eggs a female laid the larger they were.

Duckling Performance, Quality and Condition

Over the interval of 0–45 days of age, mean duckling performance differed significantly by mother category \((F_{3,37} = 2.9073, P < 0.048; \text{Fig. 4a})\): 1-NPs gained significantly less than 1-Ps, but there were no differences between 2-Ps and 2-NPs. Duckling performance did not differ during 0–15 days \((F_{3,37} = 0.0897)\), 15–30 days \((F_{3,37} = 0.41)\) or 30–45 days \((F_{3,37} = 1.7850)\).

Duckling quality (mean mass/mother) did not differ significantly at 15 days \((F_{3,37} = 0.1164)\) and 30 days \((F_{3,37} = 0.7884; \text{Fig. 4a})\). At 45 days of age, mean duckling quality varied significantly by mother category \((F_{3,37} = 3.0005, P < 0.043; \text{Fig. 4b})\): 1-NPs weighed less than 1-Ps, but there were no differences between 2-NPs and 2-Ps.

Mean duckling condition at 45 days varied significantly by mother category \((F_{3,37} = 2.8089, P = 0.05; \text{Fig. 4c})\): 1-NPs were in significantly poorer condition than 1-Ps,
but there were no differences in condition between 2-NPs and 2-Ps.

These results were not due to competitive interactions between ducklings in créches of different sizes. Ducklings/crèche had no significant functional relationship with mean duckling mass at 15 days ($r^2 = 0.004$, $N = 61$, NS) or at 30 days ($r^2 = 0.0125$, $N = 61$, NS). By 45 days, about 10% of the variance in duckling mass was functionally associated with ducklings/crèche ($r^2 = 0.105$, $N = 61$, $P < 0.01$). However, separate regressions showed that the overall relationship was due primarily to variation in mass of ducklings from 2-NPs, which was positive and significant ($r^2 = 0.22$, $N = 19$, $P < 0.043$); ducklings did better when there were more ducklings per crèche.

**Sex Allocation**

There were no mass differences by sex of eggs or ducklings at hatch, or at 15 and 30 days posthatch. By 45 days posthatch daughters weighed 103.52 ± 12.1 g less than sons (paired $t$ test: $t_{38} = -8.5904$, $P > |t| < 0.00001$). Mother category had no significant effect on the average mass differences of son and daughter eggs/mother ($F_{3,35} = 1.7453$; Fig. 5a) or on the overall percentage of sons ($F_{3,36} = 0.3970$; Fig. 5b). Mother’s mass had no significant effect on sex ratio ($r^2 = 0.0068$, $N = 37$; Fig. 5c).

**Correlations between Measures**

Egg mass was significantly positively correlated with hatching mass (Pearson correlation: $r_{39} = 0.8972$, $N = 41$, $P < 0.00001$), but not with duckling mass at 15 days ($r_{39} = 0.0199$, $P = 0.9904$), 30 days ($r_{39} = -0.0377$, $P = 0.8149$) or 45 days ($r_{39} = -0.0101$, $P = 0.9500$) posthatch. There was no correlation between egg mass and the total mass gain from hatching to 45 days posthatch ($r_{39} = -0.0337$, $P = 0.8343$). Mean hatching mass was not significantly correlated with mean duckling mass at 15, 30 or 45 days posthatch. The mean percentage of egg mass remaining at hatching was not significantly correlated with mean duckling mass at 15 days posthatch ($r_{39} = 0.2338$, $P = 0.1413$), but was at 30 days ($r_{39} = 0.4272$, $P < 0.0054$) and 45 days ($r_{39} = 0.4336$, $P < 0.0047$) posthatch. The number of ducklings/crèche was not functionally related to their mean 30 day mass ($r^2 = 0.01$, $N = 61$, NS) or mass gain from hatching to 30 days posthatch ($r^2 = 0.01$, $N = 61$, NS). There was no significant relationship between mean egg mass/mother and percentage of offspring that survived from egg to independence (offspring viability) ($r^2 = 0.000134$, $N = 39$, NS).

**DISCUSSION**

Because there were no significant differences in the mass of eggs laid by 1-NP and 1-P females, data on first-year females were silent about the CH and the MDAH. However, if males that are ‘attractive’ in groups are those preferred by individual females, data on second-year females support the CH and reject the MDAH, because second-year females laid heavier eggs for NP than P males (Fig. 3a). In addition, sex allocation, in terms of offspring sex or size of eggs by sex, did not vary by mother’s mating type (Fig. 5).

Despite a lack of significant differences in egg mass for first-year mothers, differences in the quality, performance or condition of offspring from 1-NPs emerged, and these differences could not be due to inexperience in incubation behaviour of mothers, because we incubated eggs under controlled conditions using electrical incubators. Eggs from 1-NPs used significantly more metabolic resources than eggs from 1-Ps, 2-NPs or 2-Ps (Fig. 3b), suggesting that eggs from 1-NPs experienced greater metabolic stress during incubation. Duckling mass at 45 days posthatch and mass gain from hatching to fledging were significantly lower for offspring of 1-NPs than offspring from the three other mother categories (Fig. 4).

Even though, or perhaps because, 2-NP females laid significantly heavier eggs than 2-P females (Fig. 3a), by
hatch, there were no differences in performance or quality of offspring from 2-NP and 2-P females (Fig. 3b, c). No other quality or performance differences emerged from hatching to fledging (Fig. 4). Given that the probability of survival (i.e. offspring viability) of eggs from NP pairings is significantly lower than that from P pairings (Bluhm & Gowaty 2004), increased maternal investment in eggs from NP pairings is consistent with compensation. Lack of posthatch differences in offspring between 2-NP and 2-P is consistent with successful compensation by second-year mothers with NP males.

Neither the CH nor the MDAH predicted that mother’s age would, when combined with mating type, have an effect on egg mass. Mother’s age-related variation in egg mass in response to reproduction with NP versus P males suggests that older females have greater control of resource allocation than younger females. Younger females are often less successful breeders than older female in mallards (Krapu & Doty 1979; Cowardin et al. 1985; Lokemoen et al. 1990; Dufour & Clark 2002) and other ducks (Wishart 1983; Afton 1984). Perhaps younger females in our study were developmentally unable to compensate for predictable deficits in offspring performance. The heterogeneity among mothers that we observed argues for further study of age-related variation in flexible, compensatory maternal effects. Perhaps this occurs in mallards, and if so, we would predict that eggs of older females, even those breeding for the first time, as in our study, have higher yolk androgens than eggs of younger females. If egg androgens generally affect offspring growth (Schwabl 1996), we would predict that eggs of second-year mallard mothers resulting from NP pairings contain more androgens than those resulting from P pairings.

Importantly, our results show that some mallard females have the ability to adjust some aspect(s) of egg production to allocate reserves differentially to eggs in response to reproduction with preferred and nonpreferred males. Most previous studies, as reviewed by Williams (1994), examined only the relationship between egg size and growth posthatch. However, the possible physiological mechanisms underlying such differential investment in egg production are unknown. In studies of egg composition of water birds, heavier eggs contained proportionately more yolk and less albumin (Alisauskas 1986; Hepp et al. 1986). In some species heavier eggs have higher energy density than lighter eggs (Alisauskas 1986). In precocial birds, lipid reserves increase with increasing egg size (Williams 1994). Numerous authors have suggested that females are able to adjust egg composition flexibly and adaptively (i.e. to maximize the benefits of increased egg size to offspring fitness; Alisauskas 1986). In keeping with these studies, in our study, second-year females that reproduced with NP males laid larger eggs; surviving offspring performed as well as those from P males, but overall the survivorship of eggs was significantly lower from reproduction with NP compared with P males (Bluhm & Gowaty 2004). The contributions of second-year females to the eggs of surviving offspring from NP males seemed to compensate for otherwise predictable deficits in offspring health. To our knowledge, this is the first study showing compensation for offspring viability deficits.

Our results are consistent with reproductive compensation possibly through increased maternal investment. This implies that females are allocating reserves differentially to eggs in response to mate choice or constraints on the expression of their preferences. The physiological
mechanisms underlying such differential investment in egg production are unknown. Increased maternal investment could involve differential female allocation of androgens and oestrogens and/or maternal antibodies and proteins to eggs (Schwabl 1993). The finding of female egg mass variation may have resulted from a low-level, generalized stress-induced response. It is well known that the pituitary–adrenal axis of ducks is activated in response to stress (Harvey et al. 1980). If adrenal steroids are slightly elevated, they could alter the process of vitellogenesis, a process which is heavily dependent on high circulating levels of oestrogens. Additionally, the response could be at the level of the brain or pituitary since the alteration of metabolic steroid pathways in neuroendocrine tissue in response to corticosterone may occur. This may have resulted in larger eggs being produced by females placed with nonpreferred males. For example, the eggs of NP females may have altered levels of testosterone and androstenedione, as has been found in other water birds in stressful social conditions (Reed & Vleck 2001). Importantly, the mallard females in this study continued to reproduce, so they would be a good model (see review in Bluhm 1992) for studying how maternally derived substances are regulated.

Our study was conservative in that we observed differences despite using (1) adult breeders from eggs collected in the wild and incubated and reared under controlled conditions, (2) adult virgin females for preference testing and breeding, and (3) short, 1-h preference trials in which females discriminated between only three males. However, in our study, mallard females that were paired with NP males suffered significant fitness losses (Bluhm & Gowaty 2004), including lower offspring viability (percentage of eggs that survived to fledging age) and fewer fledglings (Bluhm & Gowaty 2004) and, as we report here, significant differences in offspring performance (growth) and quality (mass at 45 days posthatch).

---

**Figure 4.** Mother category and duckling performance (a), quality (b) and condition (c). (a) Duckling mass plotted as a function of duckling age from hatching to 45 days posthatch. (b) Mean duckling mass at 45 days posthatch per mother. (c) Relative duckling condition at 45 days posthatch per mother (estimated as the residual from a regression of mean mass of offspring per mother on mean tarsus length of offspring per mother). First- and second-year females were paired with preferred (1-P, 2-P) or nonpreferred (1-NP, 2-NP) males. The dotted line indicates overall mean duckling condition.

**Figure 5.** Sex allocation of offspring based on mother category (first- or second-year female mated to preferred (P) or nonpreferred (NP) male) and mass of mother. (a) Mean difference in relative egg mass by sex (daughter – son egg mass/mother) and mother category. (b) Sex ratio (percentage of sons/mother) by mother category (using arcsine transformed data). (c) Sex ratio (percentage of sons/mother) by mother’s mass.
Although offspring viability selection probably favours mechanisms in females to reduce the likelihood of fertilization with NP males (Gowaty 1996, 1997, 2003; Gowaty & Buschhaus 1998), some females may have no other option but reproduction with NP males. In this study, as predicted by the CH, mallard mothers reproducing with NP males invested more in eggs than those reproducing with P males. However, these mothers did not produce fewer eggs, so they did not trade off the number of eggs against egg size (Fig. 1c). Our results are consistent with reproductive compensation through increased maternal investment in current offspring for second-year mothers with NP males. Compensatory investment is likely to have consequences for mothers’ future reproduction (Trivers 1972; Gowaty 2003, unpublished data; Moore et al. 2003). If compensation occurs in other species, it may explain why breeder fitness and offspring viability benefits associated with mate choice are so seldom observed.

Our results for first-year females contrast with the only other study of fitness and maternal investment in mallards. Cunningham & Russell (2000) observed that when females were paired with males that were preferred by other females (‘popular’ males), but unknown to mating females, first-year females laid significantly heavier eggs than when paired to less popular or unpopular males. No other fitness components varied with male status even though they followed duckling characteristics through to the next season. In contrast to Bluhm & Gowaty (2004), they also observed no differences in fitness components by mating type. The differences in our methods may explain the differences in the results of the two studies. Cunningham & Russell’s subjects included 16 first-year but no second-year females; ours included 23 first-year and 49 second-year virgins. They assigned males to high status if females were more likely to display towards them in mixed-sex winter flocks (i.e. they assigned male status as population-level attractiveness). We directly evaluated individual preferences free of male–male, female–female and male–female interactions that are likely to occur in mixed-sex groups (Kingett et al. 1981; McKinney 1992). They used a within-female design to evaluate sire effects. We used a between-female, parallel design to evaluate differences between females experiencing different levels of constraints, because we were concerned that matings with a given male could have residual hormonal or other carry-over effects on subsequent reproductive performance (Bluhm 1985), as might happen if females store sperm or if male seminal products affect females’ future reproduction. Cunningham & Russell (2000) reported that all 16 females in their experiment produced three clutches of eggs, whereas in our population only one of 72 females laid three clutches. Thus, the stock of mallards they used may have been from semidomesticated lineages (Cheng et al. 1980) in contrast to the wild type Manitoba mallards we used.

Despite the differences in our methods, the contrasting results of our study and Cunningham & Russell’s as well as the variation in response by mother’s age in our study argue for greater attention to variation among females. Maternal allocation of resources to adaptive, flexible maternal effects may be costly (Gowaty 2003, unpublished data). Developmental immaturity of younger females that might otherwise have avoided reproduction in their first year may have contributed to their failures to compensate. We recommend that age-related differences and other variation between females be an aspect of further study of maternal effects.

Acknowledgments

We thank Shawn Tully for his expert technical assistance in collecting egg and duckling data. We offer special thanks to Kevin Ward for advice throughout, Wyatt Anderson, Lee Drickamer, Jerry Downhower, Steve Hubbell and Allen Moore for helpful comments on the manuscript, William C. Bridges, Steve Hubbell and Wyatt Anderson for statistical advice and Randy Nelson for his encouragement. A grant from the NSF supported this work.

References


