

Seasonal decline in clutch size in European starlings: a novel randomization test to distinguish between the timing and quality hypotheses

JULIAN K. CHRISTIANS, MELISSA EVANSON and JONATHAN J. AIKEN

Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia, V5A 1S6, Canada

Summary

1. The seasonal decline in clutch size observed in many avian populations may be due to (a) a causal effect of the timing of breeding or (b) variation in individual quality (i.e. lower quality individuals lay fewer eggs and lay later in the season).
2. To distinguish between these two hypotheses, we manipulated the timing of breeding of European starlings (*Sturnus vulgaris* L.) by removing first clutches to induce the production of replacement clutches. This experiment was performed at two sites over three breeding seasons.
3. To test whether the within-individual trend in clutch size differed from the seasonal decline in clutch size observed in the population, we devised a novel randomization test that overcomes some of the potential limitations of previously used analyses.
4. We observed no within-female decline in clutch size in any year at either site. Therefore, clutch removal did not appear to manipulate female quality (due to the costs of egg production), a potentially confounding factor in such experiments; only the timing of breeding was manipulated.
5. The within-individual seasonal trend in clutch size differed significantly from the population decline in four of six cases (year–site combinations); there is good reason to believe that the lack of a significant difference in the two other cases was due to low statistical power. Thus, our data support the hypothesis that the seasonal decline in clutch size in this species is due to variation in female quality.
6. The between-year repeatability of clutch size was 0.47 and significantly different from zero, whereas that of relative laying date was much lower (–0.23) and non-significant. The repeatability data are thus consistent with the quality hypothesis: clutch size is characteristic of individuals regardless of their timing of breeding.
7. Female body mass at the beginning of incubation was not related consistently to clutch size or laying date, and therefore does not appear to be an indicator of female quality in this context.

Key-words: egg mass, intraspecific variation, laying date, Monte Carlo test; repeatability.

Journal of Animal Ecology (2001) **70**, 1080–1087

Introduction

In many avian populations, reproductive performance declines throughout the breeding season (e.g. Perrins 1970). Two classes of hypotheses have been advanced to explain such declines, although these are not mutually exclusive. The ‘timing’ hypothesis proposes that

there is a causal link between the timing of breeding and reproductive success (Verhulst, Balen & Tinbergen 1995; Hansson, Bensch & Hasselquist 2000; see also the ‘date’ hypothesis of Verboven & Verhulst 1996 and the ‘environment’ hypotheses of Winkler & Allen 1996 and Nilsson 2000). This would occur if food supply or weather conditions deteriorated throughout the season, and/or if late-breeding parents reduced their efforts due to declining chances of offspring survival (Krementz, Nichols & Hines 1989; Hochachka 1990; Svensson 1997; Lepage, Gauthier & Menu 2000; but see also Brinkhof, Cavé & Perdeck 1997). Alternatively, timing *per se* may not influence reproductive performance.

Correspondence: Julian K. Christians, Institute of Cell, Animal and Population Biology, University of Edinburgh, Ashworth Laboratories, King’s Buildings, West Mains Road, Edinburgh, EH9 3JT, UK. Phone: + 44 (0)131 6507334; Fax: + 44 (0)131 6506564; E-mail: julian.christians@ed.ac.uk

According to the quality hypothesis, the correlation between timing of reproduction and breeding success is due to a mutual relationship with a third variable, individual quality (Verhulst *et al.* 1995; Hansson *et al.* 2000; Nilsson 2000; see also the 'condition' hypothesis of Winkler & Allen 1996), such that low quality birds lay fewer eggs later in the season. Individual quality could be determined by a variety of factors such as age (e.g. Hochachka 1990), condition (e.g. Rowe, Ludwig & Schluter 1994), foraging skills or territory quality (Newton & Marquiss 1984).

One experimental technique that has been used to distinguish between these two hypotheses in avian species is the removal of first clutches; this induces relaying and thereby causes early breeding individuals to breed later in the season (e.g. Parsons 1975; Verhulst & Tinbergen 1991; Verhulst *et al.* 1995; Nilsson 2000). If the decline in breeding success were due entirely to timing, delayed individuals would be expected to show reproductive performance similar to that of unmanipulated late-nesting birds. On the other hand, if the quality hypothesis were true, then re-nesting individuals would be expected to show no decline in reproductive parameters; delayed individuals would thus have higher reproductive performance in their replacement clutch than unmanipulated birds producing their first clutch later in the season. If the reproductive performance of delayed individuals was lower than that in their first (early) clutch, but higher than that of unmanipulated late-nesting individuals, this would suggest that variation in both timing and quality play a role in the seasonal decline of reproductive performance (see Table 1). A potential problem with such experiments is that the production and care of the first clutch may have costs to the females (e.g. Heaney & Monaghan 1995; Oppliger, Christe & Richner 1996; Monaghan & Nager 1997; Monaghan, Nager & Houston 1998). As a result, the experimental manipulation may affect not only the timing of breeding, but also female quality, leading potentially to a reduction in reproductive performance in replacement clutches. Thus, such studies could exaggerate the importance of timing over quality (Hansson *et al.* 2000).

This problem will be less severe in studies of species for which the costs of egg production are relatively low, e.g. in species that complete egg production with little or no depletion of endogenous nutrient reserves. For instance, whereas laying an extra egg reduces pectoral mass in lesser black-backed gulls (*Larus fuscus* L.;

Monaghan *et al.* 1998), the depletion of endogenous protein or lipid stores is not affected by an increase in the number of eggs laid in European starlings (*Sturnus vulgaris* L.; Christians 2000). Since the nutritional costs of egg production may be relatively low in this species compared with other species, we examined the seasonal decline in clutch size in unmanipulated European starlings and in birds producing replacement clutches to test the timing and quality hypotheses. As an additional component of reproductive effort, we measured egg mass. We present a novel randomization approach for analysing the results of such experiments that overcomes some of the potential limitations of previous methods.

Materials and methods

GENERAL

Fieldwork was performed during the breeding seasons of 1996, 1997 and 1998 at two sites approximately 6 km apart near Agassiz, British Columbia, Canada (49°14' N, 121°46' W), using a nest-box population of European starlings. The experimental protocol followed the guidelines of the Canadian Committee on Animal Care (Simon Fraser University Animal Care Committee Project no. 442B; Pacific Agri-food Research Centre A.C.C. Experiment no. 9702).

Boxes were checked each morning to determine the timing of clutch initiation and completion. Eggs were measured (length and breadth ± 0.01 mm) at clutch completion and fresh egg mass was calculated from egg dimensions using an empirically derived formula. If no new egg was found on two consecutive days, the clutch was assumed to be complete. Females were captured in the nest-box after laying a first unmanipulated clutch, generally two or three days after the last egg was laid, and were weighed (± 1 g) and banded with an aluminium US Fish and Wildlife Service band to permit later identification. As part of other experiments, some females were treated with 17β -oestradiol (via silastic implants and/or intramuscular injection, $N = 37$; Christians & Williams 1999), a suspension of sheep red blood cells in saline (via intraperitoneal injection, $N = 54$; Williams *et al.* 1999) or sham versions of these treatments ($N = 38$ and 42 , respectively). Females were released and their first clutch was removed to induce the production of a replacement clutch. In addition, some birds that deserted a first unmanipulated clutch

Table 1. Interpretation of the results of clutch removal experiments with regard to the roles played by timing and quality in the seasonal decline in reproductive performance

| | Is there a difference between first and replacement clutches within individuals? | | |
|---|--|--------------------|--|
| | | Yes | No |
| Is there a difference between the within-individual decline and the population decline? | Yes | Quality and timing | Quality only |
| | No | Timing only | No seasonal decline or low statistical power |

following handling were also included in this study ($N = 25$). Following clutch removal or desertion, we continued to check nest-boxes to relocate renesting females, and to determine laying date, clutch size and egg size for replacement clutches. Females were recaptured upon completion of the replacement clutch to confirm their identity.

STATISTICAL ANALYSES

Treatment with oestradiol and sheep red blood cells had no effect on clutch size or egg mass (Christians & Williams 1999; Williams *et al.* 1999), and so all treatments were pooled for analyses. The natural seasonal trends in clutch size and egg mass in unmanipulated birds were examined using Spearman's rank correlation coefficients (CORR procedure; SAS Institute 1990); rank statistics were used since laying date (i.e. date of initiation of laying) is not normally distributed and clutch size is a discrete variable. To control statistically for variation in body mass, we used the PARTIAL statement of the CORR procedure (SAS Institute 1990); in such cases partial correlation coefficients are presented. Paired *t*-tests were used to determine whether traits differed between first and replacement clutches within individuals.

It was also necessary to test whether the within-individual seasonal trend in reproductive performance differed from the natural seasonal trend exhibited by unmanipulated first clutches (the latter set of observations will be referred to as the 'population'). Some previous studies have addressed this issue by testing whether replacement clutches differed from first clutches, controlling for date, i.e. using an ANCOVA-like analysis (Verhulst & Tinbergen 1991; Verhulst *et al.* 1995; Sanz 1999). However, there are a number of potential drawbacks to such an approach. First, this analysis requires that replacement and first clutches show the same seasonal decline (i.e. have equal slopes; Sokal & Rohlf 1995). This assumption will not necessarily be true, particularly if reproductive performance does not decline within individuals. Secondly, birds that lay a replacement clutch may not be a random sample of the population (e.g. if 'low quality' individuals do not produce a replacement clutch), and this alone could generate differences between first and replacement clutches. Finally, an ANCOVA comparing replacement clutches with the first clutches of unmanipulated birds discards data from the first clutches of renesting birds, and ignores the paired nature of the data (i.e. each renesting individual produces two clutches). If clutch size is repeatable within individuals, incorporating this paired information (e.g. the change in clutch size within individuals) could potentially reduce the variation in the within-individual trend and so increase statistical power.

To overcome these potential drawbacks, we devised a randomization method to test whether the seasonal change in clutch size within individuals differed from

the natural decline exhibited by the population. The rationale for this test was that, under the null hypothesis that the within-individual and population patterns were the same, the observed data would be a typical result of randomly drawing clutch sizes for first and replacement clutches from the same distribution. The test statistic we used to measure the extent to which the within-individual trend differed from the population trend was the difference between the within-individual decline and the population decline, where the within-individual decline was calculated as the mean of

$$\frac{(\text{clutch size of replacement} - \text{clutch size of first})}{(\text{laying date of replacement} - \text{laying date of first})}$$

among renesting birds, and the population decline was calculated as the least-squares slope from the regression of clutch size on laying date among unmanipulated first clutches.

We generated the distribution for this test statistic under the null hypothesis using a serial generalized Monte Carlo test (Manly 1997). The observed data were divided into two sets: the renesting birds and the population (i.e. unmanipulated first clutches). Observations were then swapped randomly between the two sets, one observation at a time, and the test statistic was calculated for each new permutation of the data. The following steps were used to swap observations between the two data sets: an observation from the renesting set was selected at random, and the size and date of the first clutch was swapped with a population observation, with preference given to observations with laying dates within 2 days of the renesting observation. Similarly, the size and date of the replacement clutch was also swapped with a population observation. To select a population observation, a hypothetical laying date for the replacement clutch was created by adding the laying date of the first clutch to a renesting interval (i.e. laying date of replacement – laying date of first) selected at random from the set of renesting birds. Using a hypothetical rather than actual laying date to select a population observation for swapping increased the number of population observations that could potentially be swapped with renesting observations, and so allowed more effective randomization. The randomization procedure was reiterated for a random number of steps (X), and then was performed a second time (starting again with the observed data) for Y steps, such that the total number of steps in the two runs ($X + Y$) was 999. In this way, 1000 values of the test statistic were generated (999 from random data sets and one from the observed data set) which represented the distribution of the test statistic under the null hypothesis. The proportion of the 1000 test statistics for which the absolute value was greater than or equal to the absolute value of the observed test statistic was the *P*-value of the two-sided test, i.e. the probability of obtaining data as or more deviant than the observed data, given that the null hypothesis is true. For each

test, the P -value was calculated 10 times, and the mean of these P -values is presented (Manly 1997).

Three aspects of this randomization test should be noted. First, this method calculates the change in clutch size within individuals, which overcomes the potential problem of bias that may result if the birds that renest are not a random sample of the population (e.g. a high-quality female with unusually large first and replacement clutches may have the same change in clutch size as an average female with average first and replacement clutches). Secondly, the unmanipulated first clutches of renesting birds are included in both the population and renesting data sets. Thirdly, because data are swapped between replacement clutches and unmanipulated first clutches, this approach is limited to data sets where there is substantial overlap in laying date between first and replacement clutches. Such data are desirable regardless of the method of analysis, since it eliminates the need to extrapolate the pattern of first clutches to replacement clutches.

To compare our randomization approach with the analyses used by previous authors (e.g. Verhulst & Tinbergen 1991; Verhulst *et al.* 1995; Sanz 1999), we also used ANCOVA to test whether replacement clutches differed from first clutches, controlling for date (GLM procedure, SAS Institute 1989). Unlike the randomization test, this analysis did not include first clutches from renesting birds because of the lack of independence between first and replacement clutches of the same individual.

The between-year repeatabilities of clutch size, relative laying date (number of days from the day the first egg was laid) and body mass were calculated for females which were located in more than 1 year (Lessells & Boag 1987). Randomizations and other analyses were performed using SAS (SAS Institute 1989; SAS Institute 1990).

Results

CLUTCH SIZE

Pooling data from all 3 years and both sites ($N = 473$), the modal clutch size among unmanipulated first clutches was five, with 53% of females laying this number of eggs, and the majority of other birds laying four or six eggs (16% and 23%, respectively), with a maximum of eight eggs. Approximately half (92/171) of the birds from which clutches were removed produced a replacement clutch in a nest-box. The likelihood of renesting did not appear to reflect female quality in most respects; there were no differences between females that relayed and those that did not in initial clutch size (did renest: 5.37 ± 0.08 ; did not: 5.42 ± 0.09 ; Wilcoxon $Z = 0.43$, $P > 0.2$), egg mass (did renest: 7.28 ± 0.05 g; did not: 7.20 ± 0.06 g; $F_{1,168} = 0.94$; $P > 0.2$) or body mass at clutch completion (did renest: 83.9 ± 0.4 g; did not: 84.3 ± 0.5 g; $F_{1,168} = 0.37$; $P > 0.2$). However, the probability of renesting seemed

to decline with laying date (personal observation) and for this reason clutch removal was discontinued part-way through the season.

Preliminary analyses revealed that there was significant variation in the slope of the relationship between clutch size and laying date and therefore data from the three seasons and two sites were examined separately. Figure 1 shows the within-individual and population trends in clutch size for each year and site. Examining years and sites separately, the size of unmanipulated first clutches showed a statistically significant seasonal decline in five of the six cases (Spearman's rank correlation coefficients ranged from -0.35 to -0.63 ; $P < 0.01$ in all cases). In the sixth case (site 1 in 1996), the decline was marginally non-significant ($r_s = -0.19$; $P = 0.07$). Within individuals, however, there was no decrease in clutch size between first and replacement clutches at any site in any year. At site 1 in 1997, there was actually a marginally non-significant increase in clutch size ($P = 0.07$) of almost half (0.47) an egg, or approximately 10% of the modal clutch size. In all other cases the estimated change in clutch size was less than one-fifth of an egg and was non-significant (paired t -tests, $P > 0.2$).

The lack of decline in clutch size within individual birds is consistent with the quality hypothesis (Table 1). However, the observation that the within-individual trend did not differ significantly from zero did not necessarily mean that it did differ from the population trend. Therefore, to test the timing hypothesis explicitly we examined whether the within-individual seasonal trend in clutch size differed significantly from the seasonal decline observed in the population. Using ANCOVA, replacement clutches were significantly larger than first clutches, controlling for date, in four of six cases (Table 2; the slope of the regression of clutch size on laying date did not differ between first and replacement clutches, $P > 0.2$ in all cases). Replacement clutches were not larger than first clutches at site 1 in 1996, when the decline in clutch size at the population level was relatively shallow (Table 2) and marginally non-significant (see above). The difference between replacement clutches and first clutches was also not significant at site 2 in 1998, when the number of replacement clutches was particularly low ($N = 6$). Thus, it is likely that the lack of difference between replacement clutches and first clutches in these two situations was due to low statistical power. The randomization procedure yielded similar results to the ANCOVA: the within-individual trend differed significantly from the population decline in three of the year-site combinations (Table 2). The data from site 2 in 1996 could not be analysed properly using the randomization procedure because of the poor overlap in laying date between first and replacement clutches. The consistent differences between the within-individual and population trends can be seen clearly in Fig. 1.

In four of the six cases, clutch size did not decline within individuals and the within-individual trend did

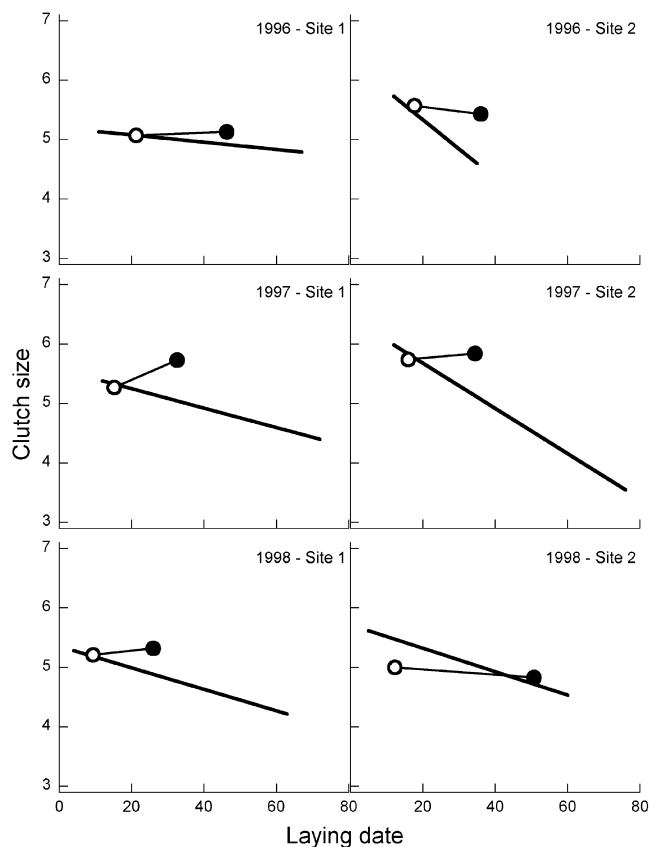


Fig. 1. Within-individual and population trends in clutch size in 1996, 1997 and 1998 at sites 1 and 2; laying date = 0 corresponds to April 1. Open and filled circles represent the means of first and replacement clutches, respectively, of females that produced a replacement clutch. The thicker solid lines show the seasonal declines in unmanipulated first clutches.

Table 2. Seasonal decline in clutch size within the population (i.e. the least-squares slope from the regression of clutch size on laying date) and within individuals (i.e. the difference in clutch size between replacement and first clutches divided by the difference in laying date between replacement and first clutches). The difference between the within-individual and population trends was tested by ANCOVA (i.e. whether replacement clutches differ from unmanipulated first clutches, controlling for date) and our randomization procedure (i.e. whether the observed data would be a typical result if the within-individual and population patterns were the same)

| Year | Site | Population | | Within individuals | | Does within-individual trend differ from population trend? | |
|------|------|------------|---------------------------------------|--------------------|---------------------------------------|--|------------------------|
| | | <i>N</i> | Decline (egg day ⁻¹) ± SE | <i>N</i> | Decline (egg day ⁻¹) ± SE | ANCOVA <i>P</i> | Randomization <i>P</i> |
| 1996 | 1 | 95 | -0.006 ± 0.004 | 15 | 0.007 ± 0.011 | 0.204 | 0.350 |
| | 2 | 28 | -0.049 ± 0.026 | 7 | 0.007 ± 0.025 | 0.039 | – |
| 1997 | 1 | 121 | -0.016 ± 0.005 | 15 | 0.033 ± 0.013 | 0.002 | 0.015 |
| | 2 | 57 | -0.038 ± 0.005 | 19 | 0.008 ± 0.013 | 0.007 | 0.045 |
| 1998 | 1 | 114 | -0.018 ± 0.004 | 34 | 0.009 ± 0.010 | 0.002 | 0.032 |
| | 2 | 58 | -0.020 ± 0.008 | 6 | -0.008 ± 0.009 | 0.803 | 0.534 |

differ significantly from the decline shown by the population, indicating that individual quality alone explains the seasonal decline in clutch size (Table 1). In the other two cases, data were consistent with the quality hypothesis, but low statistical power precluded more rigorous testing of the timing hypothesis. The between-year repeatability of clutch size was 0.47 and significantly different from zero ($P < 0.002$; $N = 35$), whereas that of relative laying date was much lower (-0.23) and non-significant ($P < 0.2$; $N = 35$). The

repeatability data are thus consistent with the quality hypothesis: clutch size is characteristic of individuals regardless of their timing of breeding.

EFFECT OF BODY MASS ON CLUTCH SIZE AND LAYING DATE

To investigate whether body mass could be a determinant of female quality, we examined the relationships between clutch size, laying date and female mass at the

Table 3. Correlations between laying date, clutch size and female body mass at the beginning of incubation among unmanipulated first clutches, and the correlation between clutch size and laying date, controlling for female body mass. Spearman's rank correlation coefficients and partial correlation coefficients are presented

| Year | Site | Correlation between body mass and | | | | Correlation between clutch size and laying date, controlling for body mass | |
|------|------|-----------------------------------|--------|-------------|-------|--|---------|
| | | Laying date | | Clutch size | | Partial r_s | P |
| | | r_s | P | r_s | P | | |
| 1996 | 1 | -0.31 | 0.0024 | 0.19 | 0.07 | -0.16 | 0.13 |
| | 2 | -0.31 | 0.11 | 0.27 | 0.18 | -0.58 | 0.002 |
| 1997 | 1 | -0.07 | >0.2 | 0.11 | >0.2 | -0.38 | <0.0001 |
| | 2 | -0.08 | >0.2 | 0.32 | 0.018 | -0.60 | <0.0001 |
| 1998 | 1 | -0.31 | 0.0013 | 0.29 | 0.002 | -0.29 | 0.003 |
| | 2 | 0.02 | >0.2 | 0.06 | >0.2 | -0.43 | 0.009 |

beginning of incubation among unmanipulated first clutches. Body mass was correlated with laying date at site 1 in 1996 and 1998, but not in 1997 or at site 2 (Table 3). The relationships between body mass and clutch size were significant in 1997 at site 2 and in 1998 at site 1, and marginally non-significant in 1996 at site 1, but not significant in the other cases (Table 3). Thus, variation in body mass did not explain variation in laying date or clutch size consistently. The seasonal decline in clutch size was still significant when controlling for body mass in five of six cases (Table 3); clutch size was not significantly correlated with laying date in 1996 at site 1 whether adjusting for body mass (Table 3) or not (see above). Therefore, body mass does not appear to be a determinant of female quality with regard to clutch size and laying date, despite being a repeatable trait (between-year repeatability = 0.60; $P < 0.0001$; $N = 34$).

EGG MASS

Mean egg mass showed no seasonal trend within first, unmanipulated clutches, and did not differ between first and replacement clutches among birds that renested; similar results were obtained whether years and sites were analysed separately or all the data were pooled (data not shown). Because there were no seasonal changes and no differences between first and replacement clutches, we did not test for differences between the population and within-individual trends. The lack of systematic changes in mean egg mass described here is consistent with the high between-year repeatability of this trait (0.76) in this population that has been reported elsewhere (Christians & Williams 2001).

Discussion

We examined whether the seasonal decline in clutch size was due to an effect of timing *per se* (either through a direct effect of timing on the ability of females to produce eggs, or to a strategic adjustment by females in response to declining chances of offspring survival), or whether clutch size and timing of breeding were both

correlated to a third trait such as female quality. Clutch size declined among unmanipulated females but not within individuals, i.e. clutch size did not differ between first and replacement clutches. Furthermore, the within-individual trend in clutch size differed significantly from the population decline in four out of six cases (year-site combinations); there is good reason to suspect that the lack of a significant difference in the two other cases was due to low statistical power (see Results). These results suggest that the seasonal decline in clutch size in European starlings is due entirely to variation in female quality (Table 1). However, body mass at the beginning of incubation was not a useful index of quality with regard to clutch size or the timing of breeding. Egg mass, another aspect of reproductive effort, showed no decline within individuals or among the population.

A potential problem with the experimental protocol used in this study is that the production of a first clutch may have costs for the females, reducing the resources available for the production of a second clutch (Hansson *et al.* 2000). However, since neither clutch size nor egg mass differed between first and replacement clutches in this study, this was not an issue; the costs of producing a first clutch would exaggerate the importance of timing (Hansson *et al.* 2000), and yet we found evidence to support the quality hypothesis only. Another potential source of bias is that our experimental protocol may have selected for higher quality birds, since not all females responded to clutch removal by producing a replacement clutch in one of our nest-boxes. However, our randomization procedure takes such potential bias into account (see Statistical analyses). Furthermore, the number and mass of eggs from the first clutch did not differ between females that did or did not reneest in our nest-boxes and it is possible that some birds produced a replacement clutch using other nesting cavities. Nevertheless, we cannot rule out the possibility that birds that reneest are of higher quality and have a different strategy with regard to clutch size than lower quality birds that do not reneest; the latter may employ a time-dependent strategy when determining clutch size.

As in this study, a number of other authors have found no decline in clutch size within individuals despite a seasonal decline among first clutches (Murphy 1986; Hochachka 1990; Wheelwright & Schultz 1994), i.e. results consistent with the quality hypothesis. Although these authors did not test explicitly for a difference between the within-individual and population trends, replacement clutches were actually larger than first clutches in one of these studies (Wheelwright & Schultz 1994) and tended to be so in another (Hochachka 1990), suggesting that the within-individual and population trends did indeed differ. In both song sparrows (*Melospiza melodia* Wilson) and savannah sparrows (*Passerculus sandwichensis* Wilson) yearling females laid smaller clutches and laid later in the season (Hochachka 1990; Wheelwright & Schultz 1994); in the case of song sparrows this was a major contributor to seasonal decline in clutch size (Hochachka 1990).

In contrast, some experiments have found support for the timing hypothesis (Parsons 1975; Batt & Prince 1979), while others have provided evidence that both timing and quality play a role in determining clutch size (Verhulst & Tinbergen 1991; Verhulst *et al.* 1995). In the case of blue tits (*Parus caeruleus* L.), one study favoured the timing hypothesis only (Sanz 1999) whereas another provided support for both hypotheses (Nilsson 2000). In addition to examining the production of multiple clutches by individuals throughout the season, a number of studies have used food supplementation to investigate the causes of the seasonal decline in clutch size. In two studies food-provisioned females produced larger clutches but showed a seasonal decline in clutch size similar to that in unprovisioned females (Nilsson 1991; Korpimäki & Wiehn 1998), suggesting that both food supply and laying date *per se* are determinants of clutch size. However, in another study supplemental food abolished the seasonal decline in clutch size (Aparicio 1994). As with the clutch removal studies, results of these studies varied within species (Aparicio 1994 vs. Korpimäki & Wiehn 1998).

An alternative approach to distinguish between the quality and timing hypotheses is to test for an effect of laying date on clutch size after statistically controlling for female quality (Winkler & Allen 1996). In our study, we examined whether female body mass at the beginning of incubation could be used as an index of female quality, but found no consistent correlation between this trait and laying date or clutch size. Body mass after egg production may not reflect quality if all females continue laying until their body condition drops to some threshold (but see Christians 2000).

To conclude, we found evidence that the seasonal decline in clutch size in European starlings was due to variation in individual quality, and not to timing *per se*. However, other such studies have provided support for the timing hypothesis, and thus the causes of the seasonal decline in clutch size appear to vary between species and perhaps even between populations and breeding seasons. Rather than test whether the timing

or quality hypotheses are correct, future work should address the conditions that determine the relative importance of variation in timing and quality to the seasonal decline in reproductive success. We hope that the randomization approach we have developed will contribute to empirical studies in this area.

Acknowledgements

We thank the staff of the Pacific Agri-food Research Centre in Agassiz, British Columbia, for permission to work on their grounds and for providing a supportive atmosphere. Evan G. Cooch provided instruction in randomization techniques and Tony D. Williams gave feedback on the manuscript as well as moral and financial support throughout this study. This research was supported by an operating grant to TDW from the Natural Sciences and Engineering Research Council of Canada (NSERC), and by NSERC postgraduate and postdoctoral scholarships to JKC.

References

- Aparicio, J.M. (1994) The seasonal decline in clutch size: an experiment with supplementary food in the kestrel, *Falco tinnunculus*. *Oikos*, **71**, 451–458.
- Batt, B.D.J. & Prince, H.H. (1979) Laying dates, clutch size and egg weight of captive Mallards. *Condor*, **81**, 35–41.
- Brinkhof, M.W.G., Cavé, A.J. & Perdeck, A.C. (1997) The seasonal decline in the first-year survival of juvenile coots: an experimental approach. *Journal of Animal Ecology*, **66**, 73–82.
- Christians, J.K. (2000) Producing extra eggs does not deplete macronutrient reserves in European Starlings (*Sturnus vulgaris*). *Journal of Avian Biology*, **31**, 312–318.
- Christians, J.K. & Williams, T.D. (1999) Effects of exogenous 17 β -estradiol on the reproductive physiology and reproductive performance of European Starlings (*Sturnus vulgaris*). *Journal of Experimental Biology*, **202**, 2679–2685.
- Christians, J.K. & Williams, T.D. (2001) Intraspecific variation in reproductive physiology and egg quality in the European Starling (*Sturnus vulgaris*). *Journal of Avian Biology*, **32**, 31–37.
- Hansson, B., Bensch, S. & Hasselquist, D. (2000) The quality and the timing hypotheses evaluated using data on Great Reed Warblers. *Oikos*, **90**, 575–581.
- Heaney, V. & Monaghan, P. (1995) A within-clutch trade-off between egg production and rearing in birds. *Proceedings of the Royal Society of London Series B*, **261**, 361–365.
- Hochachka, W. (1990) Seasonal decline in reproductive performance of Song Sparrows. *Ecology*, **71**, 1279–1288.
- Korpimäki, E. & Wiehn, J. (1998) Clutch size of kestrels: seasonal decline and experimental evidence for food limitation under fluctuating food conditions. *Oikos*, **83**, 259–272.
- Krementz, D.G., Nichols, J.D. & Hines, J.E. (1989) Postfledging survival of European Starlings. *Ecology*, **70**, 646–655.
- Lepage, D., Gauthier, G. & Menu, S. (2000) Reproductive consequences of egg-laying decisions in Snow Geese. *Journal of Animal Ecology*, **69**, 414–427.
- Lessells, C.M. & Boag, P.T. (1987) Unrepeatable repeatabilities: a common mistake. *Auk*, **104**, 116–121.
- Manly, B.F.J. (1997) *Randomization, Bootstrap and Monte Carlo Methods in Biology*. Chapman & Hall, London.
- Monaghan, P. & Nager, R.G. (1997) Why don't birds lay more eggs? *Trends in Ecology and Evolution*, **12**, 270–274.

- Monaghan, P., Nager, R.G. & Houston, D.C. (1998) The price of eggs: increased investment in egg production reduces the offspring rearing capacity of parents. *Proceedings of the Royal Society of London Series B*, **265**, 1731–1735.
- Murphy, M.T. (1986) Temporal components of reproductive variability in Eastern Kingbirds (*Tyrannus tyrannus*). *Ecology*, **67**, 1483–1492.
- Newton, I. & Marquiss, M. (1984) Seasonal trend in the breeding performance of Sparrowhawks. *Journal of Animal Ecology*, **53**, 809–829.
- Nilsson, J.-Å. (1991) Clutch size determination in the Marsh Tit (*Parus palustris*). *Ecology*, **72**, 1757–1762.
- Nilsson, J.-Å. (2000) Time-dependent reproductive decisions in the Blue Tit. *Oikos*, **88**, 351–361.
- Oppliger, A., Christe, P. & Richner, H. (1996) Clutch size and malaria resistance. *Nature*, **381**, 565.
- Parsons, J. (1975) Seasonal variation in the breeding success of the Herring Gull: an experimental approach to pre-fledging success. *Journal of Animal Ecology*, **44**, 553–573.
- Perrins, C.M. (1970) The timing of birds' breeding seasons. *Ibis*, **112**, 242–255.
- Rowe, L., Ludwig, D. & Schluter, D. (1994) Time, condition, and the seasonal decline of avian clutch size. *American Naturalist*, **143**, 698–722.
- Sanz, J.J. (1999) Seasonal variation in reproductive success and post-nuptial moult of blue tits in southern Europe: an experimental study. *Oecologia*, **121**, 377–382.
- SAS Institute (1989) *SAS/STAT User's Guide*, version 6, 4th edn, vol. 2. SAS Institute Inc, Cary, NC.
- SAS Institute (1990) *SAS Procedures Guide*, version 6, 3rd edn. SAS Institute Inc, Cary, NC.
- Sokal, R.R. & Rohlf, F.J. (1995) *Biometry*. W.H. Freeman, New York.
- Svensson, E. (1997) Natural selection on avian breeding time: causality, fecundity-dependent, and fecundity-independent selection. *Evolution*, **51**, 1276–1283.
- Verboven, N. & Verhulst, S. (1996) Seasonal variation in the incidence of double broods: the date hypothesis fits better than the quality hypothesis. *Journal of Animal Ecology*, **65**, 264–273.
- Verhulst, S., Balen, J.H.V. & Tinbergen, J.M. (1995) Seasonal decline in reproductive success of the Great Tit: variation in time or quality? *Ecology*, **76**, 2392–2403.
- Verhulst, S. & Tinbergen, J.M. (1991) Experimental evidence for a causal relationship between timing and success of reproduction in the Great Tit *Parus m. major*. *Journal of Animal Ecology*, **60**, 269–282.
- Wheelwright, N.T. & Schultz, C.B. (1994) Age and reproduction in Savannah Sparrows and Tree Swallows. *Journal of Animal Ecology*, **63**, 686–702.
- Williams, T.D., Christians, J.K., Aiken, J.J. & Evanson, M. (1999) Enhanced immune function does not depress reproductive output. *Proceedings of the Royal Society of London Series B*, **266**, 753–757.
- Winkler, D.W. & Allen, P.E. (1996) The seasonal decline in Tree Swallow clutch size: physiological constraint or strategic adjustment? *Ecology*, **77**, 922–932.

Received 9 March 2001; accepted 1 August 2001