TESTING SEXUAL SEGREGATION AND AGGREGATION:
OLD WAYS ARE BEST

CHRISTOPHE BONENFANT, JEAN-MICHEL GAILLARD, STEPHANE DRAY, ANNE LOISON, MANUELA ROYER, AND DANIEL CHESSEL

Université de Lyon, Université Lyon 1, CNRS, UMR 5558, Laboratoire de Biométrie et Biologie Evolutive, 43 Boulevard du 11 novembre 1918, Villeurbanne F-69622, France

Abstract. The study of sexual segregation has received increasing attention over the last two decades. Several hypotheses have been proposed to explain the existence of sexual segregation, such as the “predation risk hypothesis,” the “forage selection hypothesis,” and the “activity budget hypothesis.” Testing which hypothesis drives sexual segregation is hampered, however, by the lack of consensus regarding a formal measurement of sexual segregation. By using a derivation of the well-known chi-square (here called the sexual segregation and aggregation statistic [SSAS]) instead of existent segregation coefficients, we offer a reliable way to test for temporal variation in the occurrence of sexual segregation and aggregation, even in cases where a large proportion of animals are observed alone. A randomization procedure provides a test for the null hypothesis of independence of the distributions of males and females among the groups. The usefulness of SSAS in the study of sexual segregation is demonstrated with three case studies on ungulate populations belonging to species with contrasting life histories and annual grouping patterns (isard, red deer, and roe deer). The existent segregation coefficients were unreliable since, for a given value, sexual segregation could or could not occur. Similarly, the existent segregation coefficients performed badly when males and females aggregated. The new SSAS was not prone to such limitations and allowed clear conclusions regarding whether males and females segregate, aggregate, or simply mix at random applicable to all species.

Key words: Capreolus capreolus; Cervus elaphus; chi-square; isard; red deer; roe deer; Rupicapra pyrenaica; segregation coefficient; sexual aggregation; sexual segregation.

INTRODUCTION

The study of sexual segregation (Darwin 1859, 1871), the separation of males and females by habitat, spatially or socially outside of the breeding season (McCullough et al. 1989), has received increasing attention over the last two decades (Bowyer 2004). This phenomenon appears to be common among a large range of animal species (Bleich et al. 1997, Ruckstuhl and Neuhaus 2005). Several nonexclusive hypotheses have been proposed to explain sexual segregation (Bon and Campan 1996, Ruckstuhl and Kokko 2002, Bowyer 2004). Among them the “predation risk hypothesis” (Bowyer 1984, Miquelle et al. 1992, Bleich et al. 1997), the “forage selection hypothesis” (Staines et al. 1982, Clutton-Brock et al. 1987), and the “activity budget hypothesis” (Ruckstuhl 1998, Ruckstuhl and Neuhaus 2000) are the most frequently mentioned (Bowyer 2004). However, the degree to which any one of the hypotheses can account for most observed situations is still unclear (Mooring and Rominger 2004 vs. Neuhaus and Ruckstuhl 2002 vs. Yearsley and Pérez-Barbéria 2005).

Confusion about sexual segregation terminology (Bowyer et al. 1996, Main et al. 1996, Barboza and Bowyer 2000, Bowyer 2004) and a lack of a general measurement of sexual segregation hinders our ability to uncover the mechanisms driving the evolution of sexual segregation. For example, while Conradt (1998) reviewed three potential methods to quantify sexual segregation, additional measurements could have been used such as Edwards’ distance (Edwards 1971), the overlap index (D2; Manly 2005) or Nei’s distance (Nei 1972). The use of so many indices based on different metrics has prevented reliable interspecific or interpopulation comparisons despite the potential pivotal role of such comparisons to our broader understanding of the evolution of sexual segregation (Mysterud 2000).

In an attempt to develop a formal measure of segregation, Conradt (1998) proposed that the deviation of the observed group composition by sex from a random association of animals is an objective definition of sexual segregation. The resulting “segregation coefficient” (SC) provided a measure supposedly independent of stochastic variations such as sex ratio, group size, or the number of observations. It should also allow one to differentiate the relative importance of social, habitat, and spatial segregation (Bon 1991, Conradt 1998). Despite its interesting properties, the segregation...
coefficient has seldom been used (fewer than 10 reported uses out of about 100 papers published per year since 2000), likely because of its mathematical formulation and its lack of software implementation. However, the segregation coefficient (Conradt 1998) has recently been criticized as being unnecessarily conservative because it excludes “groups” composed of single animals (Bowyer 2004), which could be particularly important in some species. Thus, the segregation coefficient is unlikely to quantify adequately the annual variations in grouping patterns in slightly gregarious species. In addition, the segregation coefficient cannot quantitate the degree of segregation in the same way as it does for the degree of sexual segregation (Conradt 1998, 1999). Consequently, the segregation coefficient may not be as generally applicable as once thought so that its use is still controversial.

Here, we call for the use of the classical chi-square statistic, which can account for these problems and reliably test how sexual segregation and aggregation change over time, even when a large proportion of animals are solitary. Furthermore, we illustrate the usefulness of the chi-square statistic and reliably determine whether or not sexual segregation and aggregation occurs in three case studies of ungulate populations with contrasting seasonal grouping patterns. We suggest that coming with a measure of sexual segregation may actually be unachievable and that tests for segregation and aggregation should be preferred.

**DEVELOPING THE SSAS**

**The existing segregation coefficients**

Conradt’s (1998) index can be summarized as follows. Considering \( p_0 \) (the proportion of males in a group), \( p_1 \) (the proportion of females in a group), and \( p \) (the average sex ratio), the segregation coefficient should take on values close to 0 when males and females associate at random (all groups will tend to have \( p_0 \) nearly equal to \( p \)), and reaches a maximum of 1 where sexual segregation is complete (all groups will have \( p_0 \) equal to 0 or 1, which are at the maximum distance from \( p \)). These parameters correspond to observational data of males and females in each group in sexual segregation studies. The data include \( k \) groups with \( X_i \) males and \( Y_i \) females in the \( i \)th group that can be organized as a \( 2 \times k \) contingency table. According to Conradt (1998), the original formulation of the SC is

\[
SC = 1 - \frac{N}{X \cdot Y} \sum_{i=1}^{k} \frac{X_i Y_i}{N_i - 1}
\]

where \( N_i \) is the group size of the \( i \)th group (\( N_i = X_i + Y_i \)), \( X \) is the total number of males sampled, \( Y \) is the total number of females sampled, and \( N \) is the sum of males and females sampled. Three limitations arise from the assumptions on which SC relies. First, SC is not defined for single animals (\( N_i = 1 \)) because they are thought to segregate from their own sex and thus contribute no information about sexual segregation (Conradt 1998). However, this restriction may miss some obvious cases of sexual segregation. For instance, in the following five groups composed of 1, 1, 1, 1, 1 males and 0, 0, 0, 0, 25 females, sexual segregation obviously occurs while the SC cannot be calculated. Second, SC can be used only to investigate sexual segregation against random association and aggregation per se is not dealt with. Therefore, SC values tend to be <0, which prevent using a modified formula suggested as more appropriate (Conradt 1999). Finally, SC has no test against random association.

**Linking the segregation coefficient with the SSAS**

Testing the occurrence of sexual segregation or aggregation may be solved by considering the data as a standard contingency table and using the well-known chi-square statistic (Pearson 1900). Indeed, we can easily demonstrate that SCs are closely linked to the chi-square of the contingency table divided by \( N \) (the sexual segregation and aggregation statistic, or SSAS; Appendix A). Using the same notation as for Eq. 1, this quantity equates to

\[
SSAS = \frac{1}{N} \chi^2 = 1 - \frac{N}{XY} \sum_{i=1}^{k} \frac{X_i Y_i}{N_i}
\]

Thus, by a minor modification of the SCs towards the SSAS, one provides a general test for the segregation and aggregation patterns observed in natural populations, even when dealing with solitary animals. The SSAS being linked with Cramér’s \( V \), one of the most popular measures of association between two qualitative variables, SSAS varies between 0 and 1, regardless of the size of contingency tables (Agresti 1990). As developed in Testing the significance of segregation or aggregation, the major strength of using the SSAS is its ability to test the null hypothesis of a random association between sexes against two alternatives, segregation or aggregation. The expectancy of SSAS is \((k - 1)(N - 1)\) (Appendix B) and so is inversely related to the mean group size. Being derived from the \( \chi^2 \), SSAS provides an estimate of the distance between the observed and the expected distributions of males and females under the null hypothesis for a given number of groups \( k \) and animals \( N \). Consequently, segregation is defined as a group by sex composition that is too far from the one obtained under the null hypothesis and aggregation as a group by sex composition that is too close to the one obtained under the null hypothesis. Biologically speaking, segregation occurs when the sex ratio of each group is too different from the population sex ratio (e.g., with many unisex groups, for instance). Conversely, aggregation occurs when each group has a sex ratio almost equal to the population sex ratio.

**Testing the significance of segregation or aggregation**

The classical \( \chi^2 \) testing procedure would be appropriate for SSAS if group sizes were large. However, small
groups are common in sexual segregation studies. In such cases, a randomization procedure can be used to test the null hypothesis of independence of the distributions of males and females among groups.

We recommend the following four steps to test for sexual segregation and aggregation: (1) permute data to obtain a contingency table with the same row and column totals as the original contingency table (Patefield 1981); (2) compute the SSAS statistic for the permuted contingency table; (3) repeat steps 1 and 2 a large number of times (e.g., 999 times) to build a distribution of SSAS under the hypothesis of independence; (4) compare the observed statistic to the distribution obtained by permutation and take the appropriate statistical decision (Fig. 1a–c).

The fourth step differs depending on the alternative hypothesis. If the alternative hypothesis is sexual segregation, then using a significance level equal to $\alpha$, a $P$ value is estimated as (number of random values equal to or larger than the observed value + 1)/(number of permutations + 1). The null hypothesis is rejected if the $P$ value is less than $\alpha$. If the alternative hypothesis is sexual aggregation, the $P$ value is estimated as (number of random values equal to or less than the observed value + 1)/(number of permutations + 1). Again, the null hypothesis is rejected if the $P$ value is less than the significance level $\alpha$. Note also that a $(1 - \alpha)$ confidence interval can be computed and is bound by the $\alpha$th and $(1 - \alpha)$th quantile of the permutational distribution (Fig. 1a–c).

When multiple comparison tests are made simultaneously, a Bonferroni correction (the simplest and most conservative approach) on $\alpha$ is needed to perform an overall test at the critical value of $\alpha$. When groups are composed of solitary individuals only, all permuted tables would be equivalent (owing to the constraint that the sum of rows and columns are preserved). Here, the testing procedure would not fail but the $P$ value would equal 1 exactly, meaning that no segregation occurs. In some species, the assumption of independent animal movements among groups (Appendix A) may not be fulfilled (see also Conradt 1998:222). We suggest taking the group of associated animals instead of individual as the sampling unit.

**Material and Methods**

*The biological models*

Isard (*Rupicapra pyrenaica* Bonaparte), red deer (*Cervus elaphus* L.), and roe deer (*Capreolus capreolus* L.) have contrasting biology and behaviour (Clutton-Brock et al. 1982, Andersen et al. 1998). Red deer is a highly dimorphic species with marked segregation of the sexes. At our study site, the dressed body mass of male red deer was 30.4% larger than that of females (80.4 kg and 59.7 kg, respectively [Bonenfant et al. 2002]). A recent comparison of two red deer populations (including this population) with different timing of breeding suggested that both the predation risk and the activity budget hypotheses could explain segregation at different times of the year (Bonenfant et al. 2004). Unlike red deer, roe deer are more solitary, forming small groups only during winter (Bideau et al. 1983). Male roe deer were 7.2% larger than females in our study (average live body mass of 26.1 kg and 24.3 kg). Although very few studies have been performed on roe deer sexual segregation (but see Mysterud 1999), we expected only a limited segregation because of the slight dimorphism in size (Andersen et al 1998). The isard is a mountain ungulate, whose sexual size dimorphism is intermediate between red and roe deer (Loison et al. 1999). Despite its rather small sexual size dimorphism (average live body mass of 24.7 kg for females and 25.8 kg for males; i.e., a 6.7% dimorphism), social and habitat segregation of male and female isard have previously been documented in the Alps (Shank 1985).

*The data*

The data on the three ungulate species were collected continuously from January to December by fieldworkers.
from the Office National de la Chasse et de la Faune Sauvage (ONCFS) in France (data available online). The red deer data set includes 677 group observations made in the La Petite Pierre National Reserve (Bonenfant et al. 2002) between 1980 and 1999. The roe deer data set includes 1214 groups observed in the Trois-Fontaines “Territoire d’Etude et d’Experimentation” (Gaillard et al. 1993) in 1981 and 1982. The isard data set, including annual observations of isard in the Bazes, yields a total of 265 recorded groups (Loison et al. 2002) from 1998 to 2000. For the three study sites, group composition was observed daily either on foot or by car. Sex and age class of each individual (young of the year, yearling, or adult) was recorded. For a given month, group composition of all years were pooled as we currently have no evidence for annual changes in sexual segregation patterns (Bonenfant et al. 2004).

RESULTS

When conducting the analyses on red deer using SSAS for each month, we observed conspicuous temporal variation in sexual segregation, with significant segregation occurring only during three months (April, May, and June) and sexual aggregation occurring during three months in winter (November, January, and February; Fig. 2c). The SSAS testing procedure provided us with a more informative picture of the temporal variation in sexual segregation than the SC (Fig. 2d).

Roe deer males and females did not segregate during any month (all SSAS tests against segregation have \( P > 0.05 \)). From the tests of SSAS against aggregation, roe deer in July and December were found to show significant aggregation of the sexes (Fig. 2e). A high frequency of single animal groups (78%) was found for roe deer (mean \([±SE]\) group size, 1.27 ± 0.60). In eleven out of twelve months, the negative values of the SC suggested that male and female roe deer aggregate (Fig. 2f). However, this result is incorrect as neither segregation nor aggregation could be detected except in two months when testing with the SSAS.

In spite of the gregarious nature of isard (mean group size, 7.84 ± 7.32), aggregation of males and females did not occur in any month and sexual segregation was high and significant all year (all SSAS tests have \( P < 0.05 \); Fig. 2a). Solitary individuals were not observed in isard. Compared to SC (Fig. 2b), only the SSAS provided a clear-cut conclusion that sexual segregation in isard was significant all year-round.

DISCUSSION

Testing or measuring sexual segregation?

The use of SSAS does not lead to a measure of segregation per se, but to a test of segregation and aggregation compared to a random association of males and females. Defining a measure of dependence for categorical data is a general problem in statistics and “there may be no general solution to the problem of finding such a measure” (Lancaster 1969:239). Like the \( \chi^2 \), SSAS values cannot be used as a measure of the strength of sexual segregation (Agresti 1990). Only the test of SSAS is relevant and should be used as a decision rule (random association vs. segregation or aggregation). Comparing SSAS values is meaningless unless one estimates the confidence limits (CI) around the observed values to account for the sampling variance. However, obtaining these confidence limits assumes that the distributional properties of SSAS are known under the null and alternative hypotheses. Under the null hypothesis, \( N \times SSAS \) would follow a \( \chi^2 \) distribution if both the sample sizes are large and \( k \) (the number of groups) is held constant. Under the alternative hypothesis, \( N \times SSAS \) would follow a non-centered \( \chi^2 \) distribution at two conditions: (1) sample sizes are large and (2) an assumption detailed in Appendix A is met (Theorem 6). This second assumption cannot be verified in practice because “group” is not a fixed label and it varies in time (Appendix A, Theorem 8). This prevents estimating CI around SSAS values. Consequently, one cannot conclude that segregation is stronger or weaker in two different species or among months for a given species, by simply and directly comparing SSAS (e.g., for isard and roe deer). Nevertheless, we demonstrate below that this limitation also holds for the SC.

Several mathematical problems undermine the validity of the segregation coefficients (SC) and here we list some of the most significant. Although presented as a measure of sexual segregation (Conradt 1998), the SC is not. The expected value of SC is not 0 but is \(-1/(N - 1)\) (Appendix A, Corollary 7) in Conradt’s (1998) formulation. Similarly, using the corrected formula (Conradt 1999), the expected value of the SC still does not equal 0 and is ill-defined under the null hypothesis (the sum under the square root may take negative values; Appendix A, Corollary 7). Hence for both SC formulations, 0 cannot be taken as a baseline for a random association of the sexes. Besides, contrary to SSAS which is based on the difference between the observed and the expected distribution of males and females among groups under the null hypothesis, the SC formulation does not call for the expected values of the contingency table. Consequently, a test of SC does not warrant a test against a random association of the sexes in all cases (even with a randomization procedure). Lastly, as shown for the SSAS values, comparing two values of SCs does not provide any information about a possible difference of sexual segregation. In addition, excluding solitary animals can yield erroneous results, such as concluding wrongly to a spurious aggregation pattern, as we observed for roe deer (Fig. 2e, f). Clearly, both SC formulations cannot reliably handle aggregation patterns or be applied to only slightly gregarious species.
Even though SSAS and SC values share close mathematical formulations, their use to assess the biological patterns profoundly differs. Consequently, we call for a test based on SSAS values. The biological knowledge about sexual segregation and aggregation is better achieved with the SSAS testing procedure than with the SC. Most of time, no firm conclusion can be reached from the SC since it cannot set apart a random association of the sexes from a true segregation/aggregation process. For example, the SC value in July for the red deer is 0.20 (Fig. 2d), from which we cannot conclude whether sexual segregation occurs or not. For isard (Fig. 2b), approximately the same SC value (0.20) is found in June. Following Conradt's (1998) approach, sexual segregation should be similar in both species (same value of SC: 0.2). However, sexual segregation

**Fig. 2.** Annual patterns of sexual segregation/aggregation in (a, b) isard, (c, d) red deer, and (e, f) roe deer tested using the sexual segregation and aggregation statistic (SSAS) (a, c, e), and as given by the segregation coefficient (b, d, f). The SSAS indicates significant sexual segregation or aggregation if the observed value falls above or below the shaded area (at the 5% error level), respectively. (a, b) For isard, sexual segregation is high and significant for each month using both the segregation coefficient and the SSAS. (c, d) For red deer, significant sexual segregation occurred from April to July in addition to significant sexual aggregation in November, January, and February (c), which is similar to the pattern we found with the segregation coefficient (d). (e, f) By contrast, even though a marked seasonal pattern of group structure in roe deer is shown by SSAS (e), the segregation coefficient suggests no clear pattern over time and apparent sexual aggregation, indicated by negative values (f). Such a discrepancy arises because the segregation coefficient does not account for solitary animals (Conradt 1998).
occurs only in isard, but does not in red deer as the test of SSAS showed (Fig. 2a–c). Such discrepancies between the interpretation of SC and SSAS emphasize the importance of testing for group patterns rather than measuring any value.

Biological interpretation of SSAS

Using SSAS, we uncovered three dramatically different processes on three contrasting species. For isard, gregariousness was found all year round and associated with a year-persistent sexual segregation. Such a pattern is consistent and generally described in other mountain ungulates (bighorn sheep, Ovis canadensis Shaw, and ibex, Capra ibex L.; see Ruckstuhl and Neuhaus [2000]). Gregariousness varied according to the season for red deer, with significant segregation during three months, and significant aggregation during three months, supporting the antipredator tactic as the segregation only occurred during the calving season (Bonenfant et al. 2004). We confirmed the solitary tendency of roe deer, except during winter when they form larger groups (Bideau et al. 1983, Hewison et al. 1998). Association between male and female roe deer was mostly random, except during the mating period (July; see Bramley 1970). Hence, the three species displayed three different and contrasting grouping patterns that an average value would have not revealed.

Comparing SSAS values for a given species or among species is tempting (e.g., Bonenfant et al. 2004, Coulson et al. 2006, Loe et al. 2006), but is not recommended. Among-species comparisons with the SSAS testing procedure are possible however by building a typology of the grouping patterns. Such a classification is needed because of the strong temporal structure in group composition and group size (Bowyer 2004), and because the underlying process of sexual segregation differs according to the species’ social structure (e.g., occurrence of matrilineal groups). Using SSAS, one can distinguish sexually segregated species presenting a full segregation year-round (like the isard) from species with seasonal segregation (like the red deer).

Conclusion

The proposed SSAS approach offers biologists a general solution to the problem of how to detect both segregation and aggregation. This approach is applicable to all species at any time of their life cycle, and was already provided more than 100 years ago. We emphasize that SSAS comes with a ready-to-use function (available in the Supplement) to be run in the free software R (R Development Core Team 2006). The current SSAS testing procedure assesses the occurrence of segregation–aggregation; i.e., it does not differentiate among the habitat and the social components of sexual segregation. As SSAS is based on the $\chi^2$ theory, the decomposition of SSAS can be computed with adequate statistical tools like chi-square decompositions, log-linear models (Agresti 1990), or correspondence analyses where environmental descriptors can be entered as explanatory variables. Consequently, all these methods will allow one to separate social from habitat segregation or to investigate the effects of sex ratio or density on animal grouping behaviour.

Acknowledgments

We warmly thank everyone involved in countless hours of data collection. We greatly appreciated the contribution made by François Klein and Daniel Maillard (ONCFS) for permitting part of their data to fall into the public domain. We are indebted to Marco Festa-Bianchet, Terry Bowyer, John Lesku, Jos Milner, Atle Mysterud, Jorge A. Navarro Alberto, Kathreen Ruckstuhl, and two anonymous referees for constructive comments on previous drafts of the manuscript. The Centre National de la Recherche Scientifique (CNRS), the University of Lyon, and the Office National de la Chasse et de la Faune Sauvage (ONCFS) provided financial support of this work.

Literature Cited


Pearson, K. 1900. On the criterion that a given system deviations from the probable in the case of correlated system of variables is such that it can be reasonably supposed to have risen from random sampling. Philosophical Magazine 50:157–175.

R Development Core Team. 2006. R: A language and environment for statistical computing. R Project, Vienna, Austria.


**APPENDIX A**

Link between the sexual segregation and aggregation statistic (SSAS) and the segregation coefficient (SC) (Ecological Archives E088-196-A1).

**APPENDIX B**

Mathematical development of the SSAS (Ecological Archives E088-196-A2).

**SUPPLEMENT**

R code used to format the data and compute the SSAS (Ecological Archives E088-196-S1).