



UNIVERSITY OF CALIFORNIA PRESS  
JOURNALS + DIGITAL PUBLISHING



The American  
Ornithologists' Union

---

Unrepeatable Repeatabilities: A Common Mistake

Author(s): C. M. Lessells and Peter T. Boag

Reviewed work(s):

Source: *The Auk*, Vol. 104, No. 1 (Jan., 1987), pp. 116-121

Published by: [University of California Press](#) on behalf of the [American Ornithologists' Union](#)

Stable URL: <http://www.jstor.org/stable/4087240>

Accessed: 16/11/2012 02:20

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at  
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



University of California Press and American Ornithologists' Union are collaborating with JSTOR to digitize, preserve and extend access to *The Auk*.

<http://www.jstor.org>

# UNREPEATABLE REPEATABILITIES: A COMMON MISTAKE

C. M. LESSELLS<sup>1</sup> AND PETER T. BOAG<sup>2</sup>

<sup>1</sup>*Department of Zoology, University of Sheffield, Sheffield S10 2TN, England, and*

<sup>2</sup>*Department of Biology, Queen's University, Kingston, Ontario K7L 3N6, Canada*

**ABSTRACT.**—Repeatability is a useful tool for the population geneticist or genetical ecologist, but several papers have carried errors in its calculation. We outline the correct calculation of repeatability, point out the common mistake, show how the incorrectly calculated value relates to repeatability, and provide a method for checking published values and calculating approximate repeatability values from the *F* ratio (mean squares among groups/mean squares within groups). Received 6 February 1986, accepted 25 August 1986.

REPEATABILITY is a measure used in quantitative genetics to describe the proportion of variance in a character that occurs among rather than within individuals. Repeatability, *r*, is given by:

$$r = (V_G + V_{Eg})/V_P, \quad (1)$$

where *V<sub>G</sub>* is the genotypic variance, *V<sub>Eg</sub>* the general environmental variance, and *V<sub>P</sub>* the phenotypic variance (Falconer 1960, 1981).

In addition to its use in assessing the reliability of multiple measurements on the same individual, repeatability may be used to set an upper limit to the value of heritability (Falconer 1960, 1981) and to separate, for instance, the effects of "self" and "mate" on a character such as clutch size (van Noordwijk et al. 1980). Repeatability is therefore a useful statistic for population geneticists and genetical ecologists. Recently, we have noticed an increasing number of published papers and unpublished manuscripts in which repeatability was miscalculated. Our purpose is fivefold: (1) to outline the correct method of calculating repeatability; (2) to point out a common mistake in calculating repeatability; (3) to show how much this mistake affects values of repeatability; (4) to provide a quick way of checking published estimates, and to calculate an approximate value of repeatability from published *F* ratios and degrees of freedom; and (5) to make recommendations for authors, referees, editors, and readers to prevent the promulgation and propagation of incorrect repeatability values in the literature.

## CALCULATION OF REPEATABILITY

Repeatability is the intraclass correlation coefficient (Sokal and Rohlf 1981), which is

based on variance components derived from a one-way analysis of variance (ANOVA). The intraclass correlation coefficient is given by some statistical packages; otherwise it can be calculated from an ANOVA. ANOVA is described in most statistics textbooks (e.g. Sokal and Rohlf 1981; Kirk 1968 gives a detailed treatment of more complex designs of ANOVA), so we will not repeat it here, but give the general form of the results from such an analysis in Table 1.

Repeatability, *r*, is given by

$$r = s^2_A / (s^2 + s^2_A), \quad (2)$$

where *s<sup>2</sup><sub>A</sub>* is the among-groups variance component and *s<sup>2</sup>* is the within-group variance component. These variance components are calculated from the mean squares in the analysis of variance as:

$$s^2 = MS_W \quad (3)$$

and

$$s^2_A = (MS_A - MS_W)/n_0, \quad (4)$$

where *n<sub>0</sub>* is a coefficient related to the sample size per group in the analysis of variance. If all the sample sizes are equal (a balanced design), then *n<sub>0</sub>* is equal to the group size, *n*. If group sizes are not equal, then *n<sub>0</sub>* is smaller than the mean group size,  $\bar{n}$ . The difference between the two values increases with increasing spread in group size. The value of *n<sub>0</sub>* is calculated as:

$$n_0 = [1/(a - 1)] \cdot \left[ \sum_{i=1}^a n_i - \left( \sum_{i=1}^a n_i^2 / \sum_{i=1}^a n_i \right) \right], \quad (5)$$

where *a* is the number of groups and *n<sub>i</sub>* is the sample size in the *i*th group. If sample sizes are

TABLE 1. Analysis of variance for the calculation of repeatability.

Source of variation	df	Sum of squares	Mean squares	F ratio
Among groups	df <sub>1</sub>	SS <sub>A</sub>	MS <sub>A</sub>	F
Within groups	df <sub>2</sub>	SS <sub>W</sub>	MS <sub>W</sub>	

equal, use of this formula will yield  $n$ , so Eq. 5 may be used in all cases, including those with equal sample sizes per group. [Derivations of Eqs. 2-5 are given by Sokal and Rohlf (1981). Methods of calculating standard errors for repeatability estimates are given by Becker (1984).]

As an example, we calculated the repeatability of laying date in male and female Eurasian Sparrowhawks (*Accipiter nisus*; Newton and Marquiss 1984, Newton pers. comm.). In Table 2 we give the analyses of variance for males and females.

First we calculated  $n_0$  (Eq. 5). There were 16 males with a sample size of 2 (i.e. for which laying date was known in 2 yr), 4 with a sample size of 3, and 1 with a sample size of 5. Thus,

$$\sum_{i=1}^a n_i = (16 \times 2) + (4 \times 3) + (1 \times 5) = 49$$

and

$$\sum_{i=1}^a n_i^2 = (16 \times 2^2) + (4 \times 3^2) + (1 \times 5^2) = 125.$$

From Eq. 5,

$$n_0 = \{1/[(16 + 4 + 1) - 1]\}[49 - (125/49)] = 2.322.$$

From the analysis of variance (Table 2):

$$MS_A = 3.587$$

and

$$MS_W = 0.731.$$

Hence, from Eqs. 3 and 4

$$s^2 = 0.731$$

and

$$s^2_A = (3.587 - 0.731)/2.322 = 1.230.$$

TABLE 2. Analyses of variance of laying date in Eurasian Sparrowhawks (*Accipiter nisus*).<sup>a</sup>

Source of variation	df	Sums of squares	Mean squares	F ratio <sup>b</sup>
Males				
Among males	20	71.74	3.587	4.907***
Within males	28	20.47	0.731	
Total	48	92.20		
Females				
Among females	135	433.0	3.207	1.801***
Within females	231	411.3	1.780	
Total	366	844.2		

<sup>a</sup> Data from Newton (pers. comm.).  
<sup>b</sup> \*\*\* =  $P < 0.001$ .

Substituting into Eq. 2,

$$r = 1.230/(0.731 + 1.230) = 0.627.$$

Similarly, for females, there were 78 females with a sample size of 2, 33 with a sample size of 3, 13 with a sample size of 4, and 12 with a sample size of 5. Hence,

$$\begin{aligned} n_0 &= 2.696, \\ s^2 &= 1.780, \\ s^2_A &= 0.529, \end{aligned}$$

and

$$r = 0.229.$$

Note that the calculated value of repeatability will be negative when the  $F$  ratio is less than unity (i.e. when  $MS_A < MS_W$ ). Cases in which positive repeatabilities are reported in association with  $F$  ratios of less than one are thus a clear sign that something is amiss (e.g. Smith 1981: table 2).

A COMMON MISTAKE

Falconer (1960, 1981) referred to  $V_G$ ,  $V_{Eg}$ , and  $V_p$  as “variances” (our italics). No example of the calculation of repeatability from an analysis of variance was given in either edition of his book, although in the first edition he referred the reader to a statistical text. Sokal and Rohlf (1981: 199) stated “. . . variances are not called such in anova, but are generally called *mean squares* . . .” (their italics). Thus, it is easy to see how the

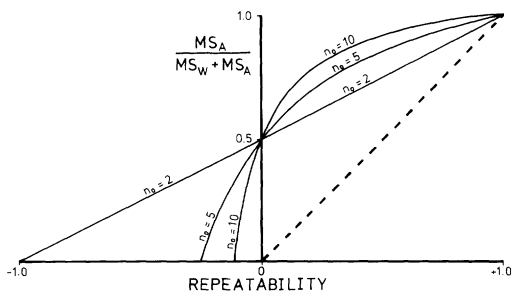


Fig. 1. The relationship between  $MS_A/(MS_W + MS_A)$  and repeatability.  $MS_A/(MS_W + MS_A) = [r(n_0 - 1) + 1]/[r(n_0 - 2) + 2]$  (Eq. A1.3). The dashed line indicates where  $MS_A/(MS_W + MS_A)$  equals repeatability.

mistake of equating Falconer's "variance" with mean square has been made frequently. A number of authors have made this error and calculated

$$MS_A/(MS_W + MS_A) \tag{6}$$

instead of repeatability (e.g. Smith 1981, Findlay and Cooke 1983, Newton and Marquiss 1984). Because

$$MS_A = s^2 + n_0s_A^2 \tag{7}$$

(Sokal and Rohlf 1981), it is obvious that  $MS_A$  not only includes the within-groups variance component ( $s^2$ ), but also depends on  $n_0$ , the coefficient representing the sample size per group. In consequence,  $MS_A/(MS_W + MS_A)$  changes systematically as the number of measurements per group increases (Fig. 1).

We cannot emphasize too strongly that the use of  $MS_A/(MS_W + MS_A)$  is not an alternative to the calculation of repeatability using variance components; it is simply wrong.

How much difference does it make?—We plotted the relationship between  $MS_A/(MS_W + MS_A)$  and repeatability (Fig. 1). This illustrated that (1)  $MS_A/(MS_W + MS_A)$  is always greater than repeatability, and in some cases considerably so (Appendix 1). For instance, when repeatability equals zero,  $MS_A/(MS_W + MS_A)$  equals 0.5. As repeatability increases to one,  $MS_A/(MS_W + MS_A)$  also increases to one, so the effect of a mistake is generally worse when repeatability is small. (2) When repeatability is between zero and one, the amount by which  $MS_A/(MS_W + MS_A)$  exceeds repeatability increases with increasing  $n_0$ . But even when  $n_0$  equals 2 the discrepancy may be considerable.

As an example, consider again the repeatability of laying date in the Eurasian Sparrowhawk. Newton and Marquiss (1984) incorrectly quoted  $MS_A/(MS_W + MS_A)$  as "repeatability." Newton (pers. comm.) supplied the data to allow us to calculate repeatability (Table 3). Comparison of repeatability with  $MS_A/(MS_W + MS_A)$  emphasizes by how much  $MS_A/(MS_W + MS_A)$  exceeds repeatability.

Checking published repeatability estimates.—An approximate value of repeatability,  $r_{approx}$ , can be calculated from the  $F$  ratio for the analysis of variance and its degrees of freedom:

$$\bar{n} = (df_1 + df_2 + 1)/(df_1 + 1) \tag{8}$$

and

$$r_{approx} = (F - 1)/(F - 1 + \bar{n}), \tag{9}$$

where  $df_1$  is the numerator degrees of freedom and  $df_2$  the denominator degrees of freedom (see Appendix 2 for derivation). If  $r_{approx}$  does not agree with the published repeatability, a check should be made whether the published value is actually

TABLE 3. Repeatability of laying date in Eurasian Sparrowhawks (*Accipiter nisus*).<sup>a</sup> Repeatability was calculated using analyses of variance and Eqs. 2–5. For further details of the analysis for all males and all females see Table 2.

	<i>F</i> ratio (df) <sup>b</sup>	Repeatability	$MS_A/(MS_W + MS_A)$
All males, irrespective of territory	4.907 (20,28)***	0.627	0.83
All females, irrespective of territory	1.801 (135,231)***	0.229	0.64
All territories, irrespective of bird	1.223 (115,210) <sup>ns</sup>	0.074	0.55
Same females on same territories	1.707 (107,108)**	0.261	0.64
Same females on different territories	1.441 (54,55) <sup>ns</sup>	0.181	0.59

<sup>a</sup> Data from Newton (pers. comm.); cf. Newton and Marquiss 1984: tables 9 and 10.  
<sup>b</sup> \*\* =  $P < 0.01$ , \*\*\* =  $P < 0.001$ , ns =  $P > 0.05$ .

TABLE 4. Repeatability of four measures of reproductive performance in female Song Sparrows (*Melospiza melodia*).<sup>a</sup>  $r_{\text{approx}}$  was calculated from the  $F$  ratios and degrees of freedom given by Smith (1981) using Eqs. 8 and 9.  $MS_A/(MS_W + MS_A)$  (quoted as "repeatability" by Smith) is given for comparison.

Character	$F$ ratio (55,91 df) <sup>b</sup>	$\bar{n}$	$r_{\text{approx}}$	$MS_A/(MS_W + MS_A)$	Correct repeatability <sup>c</sup>
No. of nesting attempts per year	1.49*	2.625	0.157	0.60	0.13
Total no. of eggs laid per year	2.55***	2.625	0.371	0.72	0.23
Total no. of young raised to 6 days of age	2.70***	2.625	0.393	0.73	0.23
Total no. of young raised to 30 days of age	0.88 <sup>ns</sup>	2.625	-0.048	0.47	-0.08

<sup>a</sup> Data from Smith (1981: table 2).  
<sup>b</sup> \* =  $P < 0.05$ , \*\* =  $P < 0.01$ , \*\*\* =  $P < 0.001$ , ns =  $P > 0.05$ .  
<sup>c</sup> Calculated using original data supplied by Smith. For these data  $n_0$  was 2.621. Several revisions to the data resulted in new  $F_{55,91}$  values. These were 1.40<sup>ns</sup>, 1.80\*\*, 1.79\*\*, and 0.81<sup>ns</sup>, respectively, giving  $MS_A/(MS_W + MS_A)$  ratios of 0.58, 0.64, 0.64, and 0.45.

$$MS_A/(MS_W + MS_A) = F/(F + 1). \quad (10)$$

If the published value is  $MS_A/(MS_W + MS_A)$ , it is incorrect. At this point there are two alternatives: (1) ask the author for sufficient information to calculate repeatability correctly, or (2) use the value of  $r_{\text{approx}}$  calculated from Eqs. 8 and 9, stating that this has been done.

As an example of this second method, we reanalyzed Smith's (1981) data on the repeatability of reproductive performance in Song Sparrows (*Melospiza melodia*) (Table 4). Smith quoted the "repeatability" of the number of nesting attempts per year as 0.60 with an associated  $F$  ratio of 1.49 (55,91 df). Substituting into Eq. 10 gives

$$MS_A/(MS_W + MS_A) = 1.49/(1 + 1.49) = 0.598,$$

approximately equal to the quoted "repeatability," revealing that this has been miscalculated. We calculated  $r_{\text{approx}}$  using Eqs. 8 and 9:

$$\bar{n} = (55 + 91 + 1)/(55 + 1) = 2.625$$

and

$$r_{\text{approx}} = (1.49 - 1)/(1.49 - 1 + 2.625) = 0.157.$$

Values of  $r_{\text{approx}}$  calculated from the  $F$  ratios are given in Table 4 for the four measures of reproductive success examined by Smith. A comparison of the values of  $r_{\text{approx}}$  and  $MS_A/(MS_W + MS_A)$  shows how much repeatability was overestimated. Smith (pers. comm.) also supplied an updated version of his original data. The final column in Table 4 gives the correct values of repeatability for table 2 in Smith (1981).

Findlay and Cooke (1983) published an ANOVA (their table 1) yielding an  $F$  ratio of 2.475 with 109, 369 degrees of freedom, and gave a "repeatability" of 0.71 for an analysis of repeatability of clutch size in Lesser Snow Geese (*Anser c. caerulescens*). Substituting into Eq. 10 gives

$$MS_A/(MS_W + MS_A) = 2.475/(1 + 2.475) = 0.712,$$

revealing that repeatability is miscalculated. Using Eqs. 8 and 9,

$$\bar{n} = (109 + 369 + 1)/(109 + 1) = 4.355$$

and

$$r_{\text{approx}} = (2.475 - 1)/(2.475 - 1 + 4.355) = 0.253,$$

one third the published value of 0.71. [Based on a slightly amended data set of 132 mother-daughter pairs, the repeatability was  $0.248 \pm 0.079$  (Findlay pers. comm.). Unfortunately, the heritability value given by Findlay and Cooke was calculated using repeatability explicitly. The correct value of heritability is  $0.208 \pm 0.071$ ,  $P < 0.01$  (Findlay pers. comm.), rather less than the published value of 0.61.]

The reason  $r_{\text{approx}}$  is only an approximate value is that  $\bar{n}$  has been used in place of  $n_0$ . If sample sizes are equal in all groups,  $\bar{n} = n_0$ , and hence  $r_{\text{approx}}$  equals repeatability. Otherwise, when sample sizes are unequal,  $\bar{n}$  is greater than  $n_0$  (Sokal and Rohlf 1981: box 9.2). As a consequence,  $r_{\text{approx}}$  is always less than repeatability (see Eq. 9); the approximation is always conservative. In our experience, the use of  $\bar{n}$  in-

TABLE 5. A comparison of  $r_{\text{approx}}$  and repeatability of laying date in Eurasian Sparrowhawks (*Accipiter nisus*).<sup>a</sup> Repeatability was calculated using analyses of variance (Table 2) and Eqs. 2-5.  $\bar{n}$  and  $r_{\text{approx}}$  were calculated using Eqs. 8 and 9, respectively.

	F ratio (df) <sup>b</sup>	$n_0$	Repeat- ability	$\bar{n}$	$r_{\text{approx}}$
All males	4.907 (20,28)***	2.322	0.627	2.333	0.626
All females	1.801 (135,231)***	2.696	0.229	2.699	0.229

<sup>a</sup> Data from Newton (pers. comm.).

<sup>b</sup> \*\*\* =  $P < 0.001$ .

stead of  $n_0$  usually does not make a large difference; as an example, we calculated  $r_{\text{approx}}$  for laying date in male and female Eurasian Sparrowhawks and compared it with repeatability (Table 5).  $\bar{n}$  is nearly equal to  $n_0$  for both males and females; in males  $r_{\text{approx}}$  differs from repeatability only in the third significant figure, and in females  $r_{\text{approx}}$  equals repeatability to three significant figures. When sample sizes differ extensively between groups, however,  $r_{\text{approx}}$  may seriously underestimate repeatability.

#### RECOMMENDATIONS

We have seen additional errors of the type discussed here in theses and unpublished manuscripts. It is ironic that this is the second time that an error in repeatability calculations has entered the ornithological literature and later been detected. The first case involved a systematic underestimation of the repeatability of egg dimensions; it took 20 years for the mistake to be corrected (Preston 1974, Kendeigh 1975).

We feel that the mistake outlined above is sufficiently common to justify the following pedagogic prescriptions for authors, referees, editors, and readers: ensure that repeatability is calculated correctly, and never publish, or use published, repeatability values unaccompanied by associated  $F$  ratios.

#### ACKNOWLEDGMENTS

Scott Findlay, Ian Newton, and Jamie Smith allowed us to publicize their mistakes in the hope of preventing others from repeating them, and also pro-

vided us with their original data to repeat the analyses. We are grateful also to the following for their comments and suggestions: Tim Birkhead, Michael Bulmer, Patrick Colgan, Fred Cooke, Bob Montgomerie, Arie van Noordwijk, Chris Perrins, Queen's TPB Group, and Mark Woolhouse. P.T.B. is supported by NSERC (Canada) grant U-0315.

#### LITERATURE CITED

- BECKER, W. A. 1984. A manual of quantitative genetics. Pullman, Washington, Academic Enterprises.
- FALCONER, D. S. 1960. Introduction to quantitative genetics, 1st ed. London, Longman.
- . 1981. Introduction to quantitative genetics, 2nd ed. London, Longman.
- FINDLAY, C. S., & F. COOKE. 1983. Genetic and environmental components of clutch size variance in a wild population of Lesser Snow Geese (*Anser caerulescens caerulescens*). *Evolution* 37: 724-734.
- KENDEIGH, S. C. 1975. Effect of parentage on egg characteristics. *Auk* 92: 163-164.
- KIRK, R. E. 1968. Experimental design: procedures for the behavioral sciences. Belmont, California, Brooks/Cole.
- NEWTON, I., & M. MARQUISS. 1984. Seasonal trend in the breeding performance of Sparrowhawks. *J. Anim. Ecol.* 53: 809-830.
- VAN NOORDWIJK, A. J., J. H. VAN BALEN, & W. SCHARLOO. 1980. Heritability of ecologically important traits in the Great Tit. *Ardea* 68: 193-203.
- PRESTON, F. W. 1974. Ancient error in a 1955 *Auk* paper. *Auk* 91: 417-418.
- SMITH, J. N. M. 1981. Does high fecundity reduce survival in Song Sparrows? *Evolution* 35: 1142-1148.
- SOKAL, R. R., & F. J. ROHLF. 1981. Biometry. San Francisco, W. H. Freeman and Co.

APPENDIX 1. The relationship between  $MS_A/(MS_W + MS_A)$  and repeatability.

By rearrangement of Eq. 2,

$$s^2 = s_A^2 \frac{(1 - r)}{r}. \quad (\text{A1.1})$$

From Eqs. 3 and 7,

$$MS_A/(MS_W + MS_A) = \frac{s^2 + n_0 s_A^2}{2s^2 + n_0 s_A^2}. \quad (\text{A1.2})$$

Substituting A1.1 into A1.2,

$$MS_A/(MS_W + MS_A) = \frac{r(n_0 - 1) + 1}{r(n_0 - 2) + 2}. \quad (\text{A1.3})$$

Subtracting Eq. 2 from A1.2,

$$\begin{aligned} MS_A/(MS_W + MS_A) - r \\ = \frac{s^2[s^2 + (n_0 - 1)s_A^2]}{(2s^2 + n_0 s_A^2)(s^2 + s_A^2)}. \end{aligned} \quad (\text{A1.4})$$

When  $s_A^2$  is positive, all terms on the right-hand side of Eq. A1.4 are positive. Hence, when repeatability is positive,  $MS_A/(MS_W + MS_A)$  is greater than repeatability. Because  $MS_A/(MS_W + MS_A)$  is always positive, it is also greater than repeatability when repeatability is negative. Thus,  $MS_A/(MS_W + MS_A)$  is always greater than repeatability.

APPENDIX 2. Calculation of  $r_{\text{approx}}$ .

$$F = MS_A/MS_W. \quad (\text{A2.1})$$

Therefore,

$$MS_W = MS_A/F. \quad (\text{A2.2})$$

Substituting A2.2 into Eqs. 3 and 4 yields

$$s^2 = MS_A/F \quad (\text{A2.3})$$

and

$$\begin{aligned} s_A^2 &= [MS_A - (MS_A/F)]/n_0 \\ &= MS_A[(F - 1)/n_0 F]. \end{aligned} \quad (\text{A2.4})$$

Substituting A2.3 and A2.4 into Eq. 2 yields

$$r = (F - 1)/(F - 1 + n_0). \quad (\text{A2.5})$$

From Eq. A2.2 it follows that

$$MS_A/(MS_W + MS_A) = F/(F + 1). \quad (\text{A2.6})$$