

# A STATISTICAL ANALYSIS OF BLACKBIRD AGGRESSIVENESS<sup>1</sup>

HAROLD E. BURTT AND MAURICE L. GILTZ

*Department of Psychology and Department of Zoology, Ohio State University, Columbus, Ohio*

## ABSTRACT

A scale with a reliability of 0.82 was developed for rating an avian behavior pattern tentatively identified as aggressiveness. The scale was applied to wild birds held in the hand and threatened in a standardized fashion. Individual differences in aggressiveness were demonstrated for Brown-headed Cowbirds and Common Grackles. Correlation between two ratings of the same bird was 0.53 for Grackles and 0.79 for Cowbirds. The most aggressive species of those studied were Grackles and Cowbirds and the least aggressive were Red-winged Blackbirds and Starlings.

## INTRODUCTION

The main purpose of this investigation was to confirm or reject the hypotheses that Brown-headed Cowbirds (*Molothrus ater ater*) and Common Grackles (*Quiscalus quiscula versicolor*) have statistically significant individual differences in aggressiveness and that the level of aggressiveness of an individual bird is fairly stable over a period of time. The data are based on a 10-point rating scale applied while a wild bird is held and threatened in a standard fashion.

The term "aggressiveness" is used advisedly. The present study is the first in a series dealing with aspects or dimensions of bird personality. This type of behavior is to be distinguished from aptitude or ability. The latter would involve how hard the bird is able to bite (in grams), while personality determines whether it *will* bite. This distinction is basic in human psychology and can be applied objectively to bird behavior without anthropomorphic implications. Terminology must be tentative until many of these aspects of personality have been explored with psychometric devices or tests. Ultimately a factor analysis of the matrix of inter-correlations of the tests will facilitate identification of the actual traits measured. For the present, "aggressiveness" appears to be the best term for the trait underlying such behaviors as biting or attacking.

## METHODS

Several methods have been used to measure aggressiveness in birds. (1) The development of peck-order is observed in an aviary or at a feeder (Ellis, 1966). (2) An experimental bird faces a single opponent in a combat cage or equivalent. (3) An experimental bird is confronted with a stuffed model (Castoro and Guhl, 1958; Lack, 1964). (4) A wild bird is held in the hand and rated on its behavior in that situation (present study).

The data were gathered in connection with a banding program at a decoy trap on the University Farm in Columbus, Ohio. The trap was of the type developed by the United States Fish and Wildlife Service (Meanley, *et al*, 1962). A rating scale was developed and was applied immediately after the bird was banded.

The bird was held in the gloved left hand with its dorsal surface toward the palm and with the bird facing upward. The right hand grasped the legs as a precaution against escape. The fingers of the left hand were opened and the bird was observed for 10 seconds; then it was threatened with the middle finger of the left hand that held it for 10 seconds; then threatened with the left thumb for 10 seconds; and finally threatened directly from the front with the middle

<sup>1</sup>Manuscript received January 24, 1968.

finger of the right hand for 10 seconds. An occasional bird attempted to fly away and had to be positioned again. This behavior represented another dimension of personality that might be investigated. The present study, however, was limited to the aggressiveness aspect.

On the basis of its behavior in response to these acts during the 40-second period of observation, the bird was rated for aggressiveness on a 10-point scale. Following a pattern used for rating people in personnel or clinical situations, the rater was guided by phrases or "bench marks" distributed along the scale. There are 10 such "bench marks", as follows:

- 10 Spontaneously begins biting rapidly and vigorously the moment the hand is opened and continues for the 40 seconds.
- 9 Spontaneously bites rapidly and vigorously, but does not begin immediately when the hand is opened.
- 8 Spontaneous, but less vigorous biting continuously through most of the initial period.
- 7 A few spontaneous bites followed by rapid and vigorous biting when threatened.
- 6 Quiet during the initial 10 seconds, but rapid and continuous biting when threatened.
- 5 Several vigorous bites at each threat, but not a continuous attack.
- 4 Some biting at each threat, but not vigorous.
- 3 No biting until the last threat, but several bites thereafter.
- 2 A single bite on the last threat.
- 1 No biting whatever; completely passive.

The rater equated the observed behavior with some particular phrase on the scale and assigned the corresponding score. For example, if the bird was quiet during the initial 10 seconds, but bit rapidly and vigorously when threatened, it was given a rating of "6".

The ratings of course are subjective in the sense that the rater made a judgment as to the correspondence between the behavior and the "bench mark." However, these subjective ratings could be checked for their statistical reliability exactly as is done with objective data, for example, the number of times a bird bites an adversary. In the present study, all the ratings were made by one person (Burtt) and their reliability was adequate (*infra*). If another rater were to use the scale, his reliability would have to be checked likewise.

#### RELIABILITY

The value of the data obtained by the foregoing methods depends on the reliability of the rating scale. With an instrument such as an aptitude test, it is standard practice to administer the test twice in succession to the same group of people and compute the coefficient of correlation between these pairs of scores. The maximum possible coefficient is 1.00 and the more closely the computed coefficient approximates this, the higher the reliability of the instrument.

With the aggressiveness scale, it was necessary to modify the procedure so that the second rating would not be influenced by knowledge of the first rating of the same bird. The reliability was determined for a sample of 100 Cowbirds and 100 Grackles. Each of these birds was taken from the gathering cage, banded, rated, and then put into a second gathering cage nearby. Birds of other species in the first gathering cage were released immediately after banding. When the first cage had been emptied, the birds in the second cage were rated again without consulting the first ratings. The interval between ratings usually was about 15 minutes. The probability of environmental changes altering the bird's personality during that interval was small; prior to both ratings, the bird was in a cage 24 x 18 x 12 inches with other birds.

In this way two ratings of aggressiveness were available for each of the 200 birds. The correlation between the two ratings was computed. The product-

moment correlation  $r$  (Edwards, 1951, p 79) was used throughout the study. In the present case,  $r$  is 0.82. This reliability compares favorably with that of rating scales at the human level and justifies the use of the present technique.

#### INDIVIDUAL DIFFERENCES

Individual differences may be evaluated by using "multiple repeaters," i.e. birds which have been re-trapped several times and rated for aggressiveness on each occasion. There are 22 Grackles on which five to 16 ratings were made. With these cases a simple analysis of variance will confirm or reject the null hypothesis that the birds do not differ in aggressiveness. Table 1 shows that the ratio (F) of the variance between birds to the variance within birds (Edwards, 1951, p. 198) is 4.4. Entering a table of values of F with the appropriate degrees of freedom, it is found that a ratio of 2.0 is sufficient for significance at the 1% level. Thus the data reject the null hypothesis and the Grackles differ significantly ( $p < .01$ ) in aggressiveness.

TABLE 1

*Individual differences as determined by simple analysis of variance for birds trapped repeatedly*

	F-ratio of variance <i>between</i> birds to variance <i>within</i> birds	F-ratio needed for significance of individual differences at the 1% level
Grackles	4.4	2.0
Cowbirds	21.6	2.2

Another approach to these individual differences is possible. The aggressiveness scores for a given bird may be averaged (mean) and the standard deviation computed. The scores for a second bird may be treated likewise. Then the difference between the two means may be evaluated by a t-test (Edwards, 1951, p. 176) to determine its significance. In this way, each mean in the sample may be compared with every other mean to determine how many of the differences are significant. However, the analysis of variance (*supra*) establishes in essentially one operation, and thus more expeditiously, the existence of significant individual differences.

Among the Cowbirds, there were 17 for which five to 13 ratings were made. The F-ratio (Table 1) is 21.6, while a ratio of 2.2 is sufficient for significance at the 1% level. Obviously the null hypothesis is rejected and, like the Grackles, the Cowbirds have significant individual differences in aggressiveness ( $p < .01$ ). Starlings (*Sturnus vulgaris vulgaris*) and Red-winged Blackbirds (*Agelaius phoeniceus phoeniceus*) were trapped in considerable numbers but did not "repeat" sufficiently to implement an analysis of variance like the foregoing.

#### STABILITY

The question may be raised as to the stability or consistency of a bird's aggressiveness, i. e. the extent to which its behavior in this respect remains at about the same level over a period of time. In the preceding analysis (Table 1), the fact that the variance within birds is less than the variance between birds implies that an individual bird is stable in aggressiveness. However, a more direct approach to the matter is to consider all the birds of a given species for which there are available two ratings separated by an interval of at least one day and to correlate the two ratings. This procedure should be distinguished from the correlation between two ratings made in immediate succession for determination of reliability

(supra). In the latter case, it was assumed that the bird's personality would not change in a few minutes and thus the correlation coefficient really indicated whether the rater changed. In the former case, it is a question of whether the personality of the bird changes during the longer interval.

For double ratings on 95 Grackles, separated by at least the interval of one day, the correlation ( $r$ ) between the two ratings is 0.53. A t-test rejects ( $p < .01$ ) the hypothesis that the true correlation is zero. Similarly, with 132 Cowbirds,  $r$  is 0.79 and is likewise significant ( $p < .01$ ). There is no significant sex difference in stability for the Cowbirds. For 106 males  $r$  is 0.78 and for 26 females 0.83. Complete data on the sex of the Grackles are not available.

The variation in the length of time between ratings does not appear to introduce any error. The  $r$  based on ratings separated by 5 days or less and the  $r$  based on ratings separated by more than 5 days do not differ significantly ( $p > .05$ ). This is true for both species. The conclusion appears warranted that aggressiveness as measured in the present study is a stable aspect of a bird's personality.

#### SPECIES DIFFERENCES IN AGGRESSIVENESS

Other species that were banded in considerable numbers were also studied, though less intensively. Single ratings were obtained for 400 Red-winged Blackbirds, 400 Starlings, and 100 Mourning Doves (*Zenaidura macroura*). When every one of the 100 doves received a rating of 1 on the aggressiveness scale, there seemed no point in enlarging the sample. Table 2 gives the means and standard deviations for the ratings of these three species, as well as for the initial ratings of all the Grackles and Cowbirds that were handled.

TABLE 2

*Species differences in aggressiveness*

	Mean Score	Standard deviation of the distribution of species	Number of birds in sample on which values are based
Mourning dove	1.0	0	100
Starling	1.64	1.03	400
Redwing	1.80	1.32	400
Cowbird	2.96	2.39	831
Grackle	2.95	2.14	911

In evaluating the significance of the data, the frequency distribution curves for the species were first compared, because all the curves, except that for Mourning Doves, were skewed toward the smaller aggressiveness scores. Chi-square tests were run on the distributions for all combinations of two species to check the null hypothesis that the distributions were drawn from the same sample of aggressiveness (Edwards, 1951, p. 239). The "expected" values in the computation were obtained by splitting the total frequency of the two species in a given class interval according to the percent which each species constituted of the entire sample. If the hypothesis was not rejected at the 5% level or better, nothing further was done. If, however, the hypothesis could be rejected, then a t-test was applied to the difference between the means of the two species. On this basis, Cowbirds do not differ significantly from Grackles, nor do Starlings differ significantly from Red-winged Blackbirds. For all the other combinations, the distributions differ significantly ( $p < .01$ ) and the means differ significantly ( $p < .01$ ). No significant sex differences in aggressiveness are found between distributions or between means in the Cowbirds or in the Red-winged Blackbirds. In general, of the

species studied, Mourning Doves are the least aggressive, Starlings and Red-winged Blackbirds are somewhat more aggressive, and Cowbirds and Grackles are the most aggressive of all.

It is possible that aggressiveness may vary somewhat during the year because of physiological changes often associated with mating, nesting, and migration. The available sub-samples are inadequate to yield statistically significant results on this point.

#### DISCUSSION

While rating scales have seldom been used in ornithology, they have been accepted instruments in psychological personnel work for 50 years (Burtt 1942, Chapter 12; Uhrbrock 1950). With any psychometric device, there is always a moot question as to just what it is measuring. A bird held in the hand is probably frightened to some extent. While an aggressive manifestation, such as biting, cannot be an actual expression of fear, it may nevertheless be modified or partly inhibited by the fear. Much the same thing is true in a peck-order situation, or even in a combat cage where fear of the opponent may be present. In all these situations, there may be degrees of fear (timidity), just as there are degrees of aggressiveness. It is possible that what is involved is a personality continuum with extreme timidity at one end, extreme aggressiveness at the other and various mixtures in between. However, observations in the trap and in the testing situation have led the writers to believe that the two traits are orthogonal and that the behavior in the hand, as sampled by the rating scale, stems from an underlying personality trait called, for the present, "aggressiveness". Another study (unpublished) developed a technique for measuring "complacency," a trait empirically orthogonal to the aggressiveness variable. Correlations between the two average about 0.10 for several species.

As suggested at the outset, the detailed structure of bird personality must await data from a variety of psychometric techniques and factor analysis of the inter-correlations. However, for the present, bird behavior has been evaluated, in the context of personality, with a scale that has adequate statistical reliability and is expeditious in administration. It would seem profitable to compare such measurements with other behavioral variables.

#### SUMMARY

1. A scale was developed for rating the aggressiveness of a bird held in the hand and threatened in a standard fashion.
2. The scale has a reliability of 0.82.
3. Statistically significant individual differences in aggressiveness were demonstrated by a simple analysis of variance for Cowbirds and Grackles.
4. Rating the same bird, when it is trapped on two separate occasions more than a day apart, gives an  $r$  between the two ratings of 0.53 for Grackles and 0.79 for Cowbirds.
5. Grackles and Cowbirds prove to be the most aggressive species in this study. Red-winged Blackbirds and Starlings are significantly less aggressive, and Mourning doves give no indication of aggressiveness.

#### LITERATURE CITED

- Burtt, H. E. 1942. *Employment Psychology*, New York, Harpers. 568 p.
- Castoro, P. and A. M. Guhl. 1958. Pairing Behavior of Pigeons Related to Aggressiveness and Territory. *Wilson Bull.* 70: 57-69.
- Edwards, Allen L. 1951. *Statistical Analysis for Students in Psychology and Education*, New York, Rinehart. 360 p.
- Ellis, C. R., Jr. 1966. Agonistic Behavior in the Male Starling. *Wilson Bull.* 78: 208-224.
- Lack, D. 1964. *The Life of the Robin*. Second Edition, London Witherby. Ch. 12.
- Meanley, B., J. S. Webb, D. P. Fankhauser, and R. T. Mitchell. 1962. The Blackbird Banding Program of the Patuxent Wildlife Research Center. *Eastern Bird Banding Association News* 25: 117-121.
- Uhrbrock, R. S. 1950. Standardization of 724 Rating Scale Statements. *Personnel Psychology* 3: 285-316.