It is clear from these observations that during the breeding season, the behavior of a Gila Woodpecker depends to a large extent on what its mate is doing. So while the pronounced size dimorphism disposes the sexes generally to perform different parental functions, these roles are not exclusive; they can be traded back and forth between mates.

Effect of litter on leaf-scratching in emberizines.— Many species of emberizines turn leaves and other litter by a two-footed scratching movement resembling hopping, in which the litter is thrown rearward under the bird (Hailman, Wilson Bull. 85:348–350, 1973). These scratches are performed sequentially in bouts, where the probability of adding another scratch to a bout is constant and hence independent of the number of scratches already performed in the bout (Hailman, Wilson Bull. 86:296–298, 1974). The quantitative model expressing this relationship predicts that the log frequency of bouts having s or more scratches ($\log f_s$) is a linear function of the number of scratches/bout (s):

$$\log f_s = (s - 1)\log p + \log B, \qquad (1)$$

where p is the constant probability of adding another scratch, $\log p$ is the slope and $\log B$ is the intercept of the linear regression. The present study evaluates one ecological variable previously suggested as possibly affecting the value of p: the amount of litter on the ground.

Equation (1) predicts quantitatively the behavior of several species: the White-throated Sparrow (*Zonotrichia albicollis*) and Dark-eyed Junco (*Junco hyemalis*) originally studied (Hailman 1974) and further considered in the present study, the White-crowned (*Z. leucophrys*) and Fox (*Passerella iliaca*) sparrows studied subsequently (Hailman, Wilson Bull. 88:354–356, 1976), and Rufous-sided Towhee (*Pipilo erythrophthalmus*) studied independently by E. H. Burtt, Jr., and me (Burtt and Hailman, Wilson Bull. 91:123–126, 1979). Furthermore, Burtt showed that p depends in part upon the amount of food available, in that towhees scratch in longer bouts as food becomes scarcer. I had suggested that p depends in part on the amount of litter (Hailman 1974), so it is possible that p is a compound variable, and the present experiments were set up to test the effect of litter.

The study plot consisted of a rectangle outside my study window in Madison, Dane Co., Wisconsin. The plot was divided in half, creating north and south meter-square quadrats. All litter was raked to the dividing line between the two quadrats, one measuring cup of about 235 cm³ of mixed bird seed was scattered homogeneously over each area, and then all litter was raked over one of the two plots, thus creating a "littered" and a "bare" area. Wind, squirrels, and the birds themselves quickly scatter litter so that the bare area does not remain truly bare for long, nor does the littered area remain homogeneously littered; there is, however always a distinct difference in the amount of litter on the two sides.

I observed white-throats and juncos foraging the two areas for about 2 days, then reraked

and reseeded the plots, and switched sides by raking the litter onto the previously bare plot. The first littered plot was chosen at random, and then alternated with each reseeding. The study was conducted in autumn of 1973–76 and 1979, between the time that leaves had mostly fallen and the first snow covered the ground. The study is based on 1584 bouts containing a total of 2316 scratches observed during 39 observation periods on 25 different dates.

Statistical treatment of the data consists of least-squares regression of log f_s and s, with the goodness-of-fit to linear relationship expressed by the coefficient of determination (r^2). This statistic, which is the square of the correlation coefficient (r), more sensitively indicates the degree of fit than a simple test for the statistical significance of linearity, which is always highly significant in these data. The central tendency could be expressed by the mean number of scratches/bout, but a more sensitive measure is the probability (p) of adding another scratch—the antilog of the slope in equation (1). There is apparently no completely valid statistical technique for evaluating the difference between two regression lines of cumulative frequencies (indeed, past contributions in this series have attempted no such tests).

In ordinary linear regression, the slope-coefficient a (=log p) approaches a normal deviant whose variance (V) is the square of the slope's calculated standard error (equations found in most texts on parametric statistics), so I have evaluated a difference in slopes as

$$t = (a_1 - a_2)/(V_1/N_1 + V_2/N_2)^{1/2}, \tag{2}$$

where N is the number of points on the regression line for a data set. The number of degrees of freedom (df) is only approximately calculatable in such a t-test using unknown variances that cannot be assumed equal and the calculation itself is arduous, so I have used simply df = $N_1 + N_2 - 2$, which is strictly valid only when the variances are assumed equal. In all cases I have tested for equality of variances using Fisher's ratio (F = larger V/smaller V) and the variances could not be shown significantly different in any case where a t-test indicated a significant difference in slope. These methods are not strictly valid because f_s at a given s is not independent of f_s at some other value of s, so the reader may accept or ignore the outcome of t-tests as he pleases: it is a crooked wheel, but the only one in town.

The first concern was whether there is some systematic bias in the north and south plots. I recorded B = 306 bouts of scratching by white-throats in the north and 234 bouts in the south plot when littered, and the calculated probabilities of adding an additional scratch were found to be p = 0.30 and 0.23, respectively; the difference in slopes was not significant (t = 1.354, df = 7, P_t = 0.22, two-tailed). Similar data from the junco yielded p = 0.31 and 0.32 for 301 and 291 bouts, respectively, with an insignificant difference (t = 0.203, df = 10, P_t = 0.84, two-tailed). When the plots were bare for the white-throat, the values were p = 0.14 and 0.10 (t = 1.30, df = 4, P_t = 0.26, two-tailed) for 48 and 198 bouts, and for the junco were p = 0.28 and 0.09 for 122 and 76 bouts (sample too small for statistical test because V cannot be calculated in the latter regression). Therefore, no differences between the two plots could be established, and the data from the two were combined for testing between littered and bare areas.

Fig. 1 shows that, as predicted, a greater amount of litter causes White-throated Sparrows to scratch in longer bouts. The variances are not different (F = 1.88; df = 2,4; P_F = 0.265), but the difference in slopes is highly significant (t = 5.16, df = 6, P_t = 0.0011, one-tailed). Similar data for the Dark-eyed Junco (Fig. 2) show the same pattern, but the difference in slope is not nearly so great. The F-ratio shows the variances to be significantly different (F = 8.78; df = 4,5; P_F = 0.17), so t-tests were run both with df = $N_1 + N_2 - 2$ and the more arduously calculated approximation of equation (2). In both cases the difference in slopes was not quite statistically significant (t = 1.60, P_t's = 0.072 and 0.085, one-tailed).
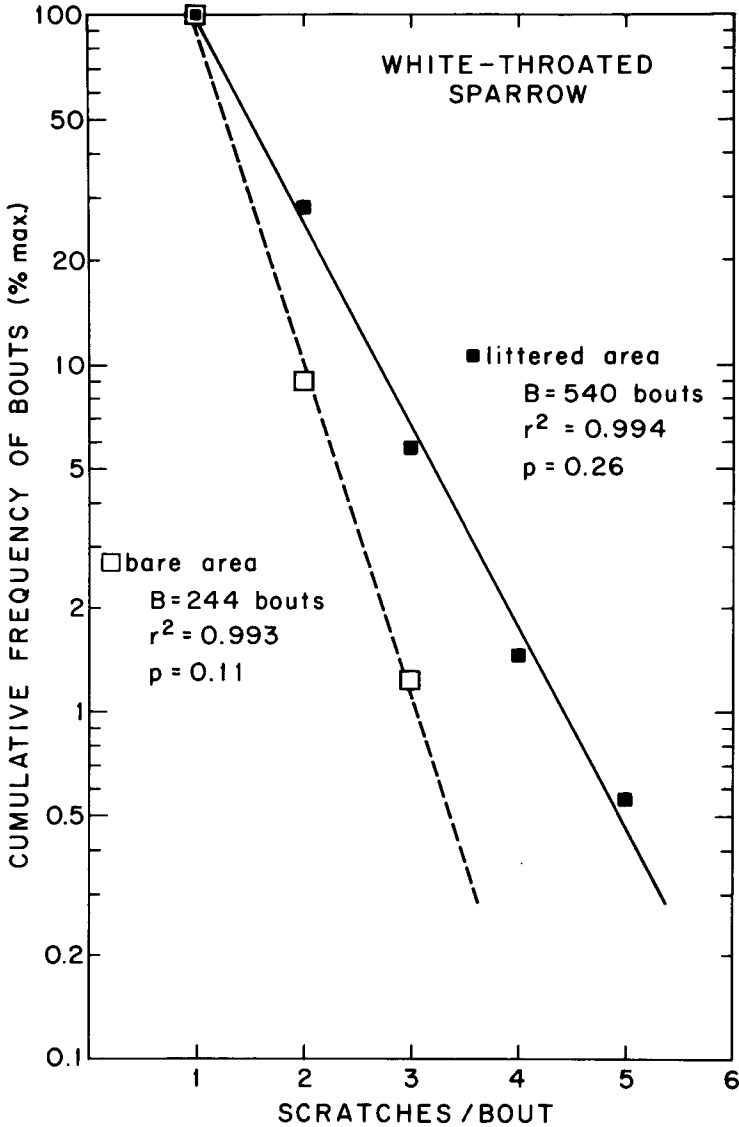
FIG. 1.   Linear relationships between s (scratches/bout) and the frequency of bouts having s or more scratches plotted logarithmically (log f_s) for the White-throated Sparrow. B is the number of bouts of scratching observed and r^2 the coefficient of determination expressing the fit to linearity. The lines drawn are fitted by least-squares regression. The difference in slopes between littered and bare areas is statistically significant (see text).
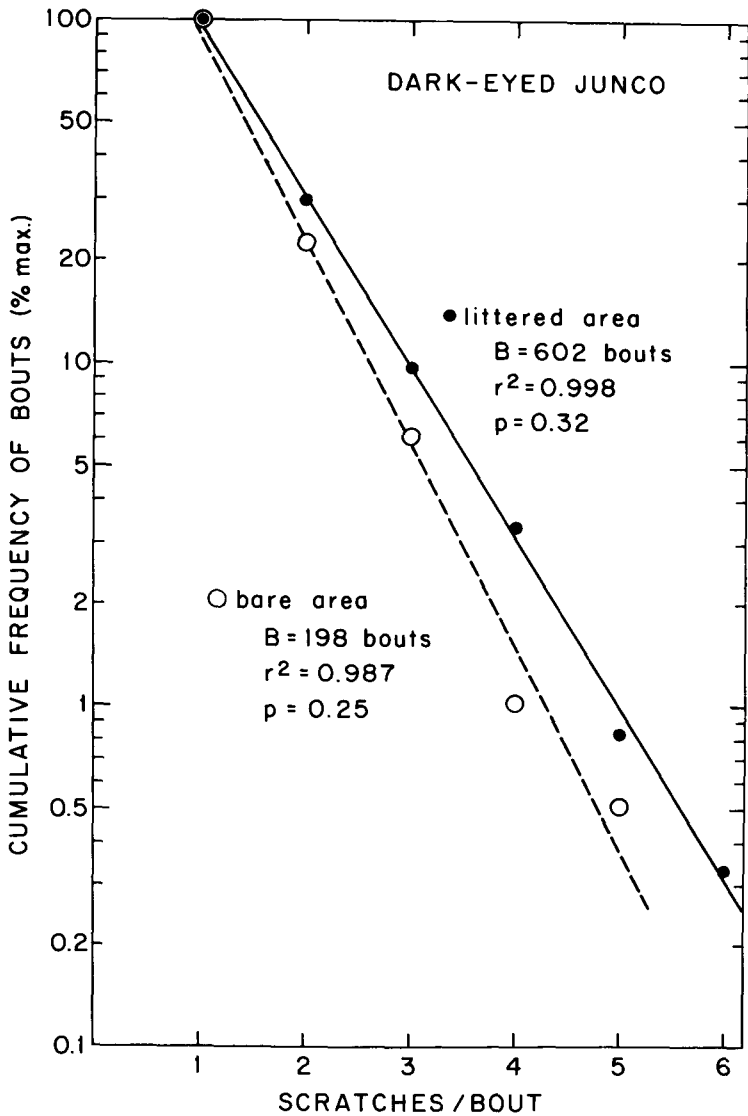
FIG. 2.   Similar linear relationships in the Dark-eyed Junco (see Fig. 1). The difference in slopes is not quite statistically significant (see text). For evaluation of differences between the two species in a given area, see statistical comparisons given in the text.

Therefore, even though the difference is in the predicted direction of longer scratching bouts in the leaved area, the difference cannot unequivocally be established as real.

When I originally looked at these two species, I had the impression that white-throats scratched in longer bouts than did juncos, and this initial impression was consistent with slight differences in slopes of their data when later studied (Hailman 1974). However, I noted that the difference might be due to differences in scratching sites chosen by the two species, white-throats choosing more littered sites. The present data allow species comparisons within a particular site. In both sites it is the junco, not the white-throat, that has an absolutely higher probability of adding scratches to a bout (compare p-values of Figs. 1 and 2). In the littered site this difference is not significant ($t = 1.96$, df = 9, $P_t = 0.082$) but it can be established as real in the bare site ($t = 4.22$, df = 6, $P_t = 0.006$). It appears, then, that juncos rather than white-throats scratch in longer bouts under similar ecological conditions.

From the results in the figures, I conclude that the amount of litter is a variable affecting the length of scratching bouts. It should be pointed out that there was no control for food in these experiments, begun years before Burtt's experiments showing the importance of food in determining the length of bouts. At the beginning of each reseeded plot the amount of food was equal, but is seems likely that in the bare area seed became scarcer more rapidly than in the littered area, thus raising the value of p in the bare area and hence minimizing the differences in slope shown in the figures. Despite this mitigating effect, the present experiments established the reality of litter as another factor determining the probability of scratching. It is not possible on the basis of present evidence, however, to combine quantitatively the effects of litter and food on scratching; we can say only that both play a role in the quantitative determination of emberizine foraging. Nor do litter and food exhaust the possible variables affecting the probability of scratching; the hunger of the bird, for example, may also play a role, as may social factors such as companions scratching nearby.

The apparent contradiction between this and the previous study (Hailman 1974) with regard to species' differences is probably attributable to a number of factors. First, the difference found in the previous study is small and probably trivial. If real, it may have been due to observing white-throats more frequently in heavily leaved areas where the probability of scratching is relatively high, and juncos in less littered areas where the probability of scratching is lower. The present study was not constructed so as to measure preference for foraging site and any possible difference in site-preference cannot be resolved by data presently available.

Finally, it seems worth pointing out that these experiments serve to emphasize the importance of quantitative testing of intuitive hypotheses derived from anecdotal observations. The results in one case confirm the incidental observations that litter helps determine the length of scratching bouts, but in the other case reject the incidental observations that white-throats scratch in longer bouts than juncos. We thus come one step closer to understanding in detail the foraging behavior of a species and the nature of differences among species.