significant effect. However, the results do show that ECS had a disrupting effect if given directly after the five learning trials.

Results of this study support the consolidation hypothesis in showing that retroactive amnesia occurs without aversive effects when shock is not given repetitively and that time of administration affects retroactive amnesia, while location does not. These results are consistent with those of Hudspeth, McGaugh, and Thompson (*7*), who have shown that aversive effects produced by ordinary shock become more intense over trials than aversive effects produced by ECS. They also found that aversive effects produced after repeated ECS sequences were stronger than those produced by the first sequence. The results of the present study suggest that the findings of Lewis and Adams (*4*) were probably due to the aversive effects of repeated ECS sequences. Thus, avoidance and competing response explanations are not adequate to handle retroactive amnesia produced in a single-ECS situation.

We are of the opinion that perhaps the single-shock and the sequential-shock techniques are each appropriate to answer different questions (for example, the relation between ECS and retroactive amnesia on the one hand and the relationship between the shock and aversion on the other). Thus, it is not a matter of which technique or theory is or is not supported; rather it is a matter of which technique is more fruitful or appropriate to answering a particular question.

DWIGHT J. LEONARD*
ALBERT ZAVALA†
Department of Psychology, Kansas State University, Manhattan

References and Notes

1. N. E. Miller and E. E. Coons, *Amer. Psychologist* **12**, 394 (1955), abstr.
2. E. E. Coons and N. E. Miller, *J. Comp. Physiol. Psychol.* **53**, 524 (1960).
3. H. E. Adams and D. J. Lewis, *ibid.* **55**, 299 (1962); *ibid.*, p. 302.
4. D. J. Lewis and H. E. Adams, *Science* **141**, 516 (1963).
5. C. A. Pearlman, S. K. Sharpless, M. E. Jarvik, *J. Comp. Physiol. Psychol.* **54**, 109 (1961).
6. M. C. Madsen and J. L. McGaugh, *ibid.*, p. 522.
7. W. J. Hudspeth, J. L. McGaugh, C. W. Thompson, *ibid.* **57**, 61 (1964).
8. Paper was presented at the Midwestern Psychological Association, St. Louis, May 1964, by one of us (D.J.L.).

* Present address: U.S. Army Training Center, Human Research Unit, P.O. Box 787, Presidio of Monterey, Calif.
† Present address: American Institutes for Research, 1808 Adams Mill Road, NW, Washington, D.C.

19 August 1964

Statistical Models for Predicting Numbers of Plant Species

In "Species abundance: natural regulation of insular variation" [*Science* **142**, 1575 (1963)], Hamilton, Rubinoff, Barth, and Bush have compared two statistical models for predicting numbers of species from environmental factors. The first is a "linear" model, $y = bx$. The second is a "curvilinear" model $y = bx^z$ where $z \neq 1$. Although they give no exact statement of their test of "goodness of fit," they conclude that model 1 is the superior.

Considering only the relation of species numbers (Y) to area (X_1), much of the question of the form of the relation can be determined by a quick look at a simple scatter diagram. Figure 1 shows a "natural" and Fig. 2 a "logarithmic" plot of the data. Also shown on each plot is the least squares fitted line for a simple regression.

It is readily apparent from Fig. 1 and 2 that the "scatter" about the line of Fig. 2 is more regular than that for Fig. 1. One of the assumptions of the use of goodness-of-fit tests is that the distribution of errors about a fitted line be independent of the value of the independent variable. This condition is not met for model 1.

If errors are to be transformed for comparison, they must be transformed in a manner that is consistent with the regression equation used. Therefore, the sums of the squared deviations

$$\Sigma (Y_1 - \bar{Y})^2$$

and

$$\Sigma (Y_2 - \bar{Y})^2$$

are not directly comparable. Perhaps one might compare the two errors by converting the error in log units at the mean value of the relation of Fig. 2. This, then, would be grossly comparable with the standard error of estimate from Fig. 1. For a mean of 119.3 and a standard error of 93.7 species, the range of the confidence interval at the mean is from 26 to 213 species for model 1. The comparable range for model 2, with a standard error of 0.331 log units, is 56–255. The exclusion of one extreme value, Albemarle, changes the equation for model 1 from

$$y = 95.4 + .12X \text{ (SE 95)}$$

to

$$y = 70.3 + .48X \text{ (SE 86)}.$$

The comparable change for model 2 is from

$$y = 28.6\ X^{0.331} \text{ (SE 0.319)}$$

to

$$y = 28.2\ X^{0.339} \text{ (SE 0.332)}.$$

Figure 1 also gives a good visual picture of why "Model 1 predicts floral richness for larger islands more accurately than it does for smaller islands." This is a result of the least squares fitting itself. The few larger islands receive much more weight than all the data for smaller islands. The two figures give an indication of why X_1 is significant for model 2 but not for model 1.

The expected relation of errors to the "true" values may give some perspective to the problem of choosing a model. Assume an island were observed to have 10 species and a "satisfactory guess" would be in the range of 5 to 15 species, or ± 5 species. If a second island were observed to have 1000 species, would a "satisfactory guess" need to fall between 995 and 1005 species? This is the requirement of model 1. Probably, the proportional errors of model 2 would be preferable. If a guess of 6 to 17 is good enough for 10, then a guess of 600 to 1700

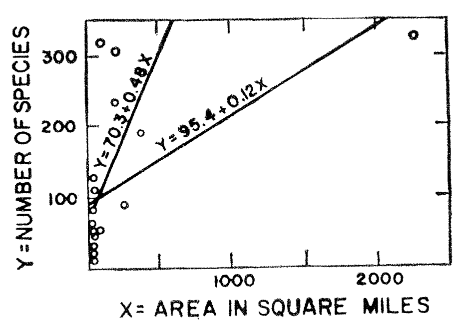

Fig. 1. Relation of species number to area, plotted to natural scales.

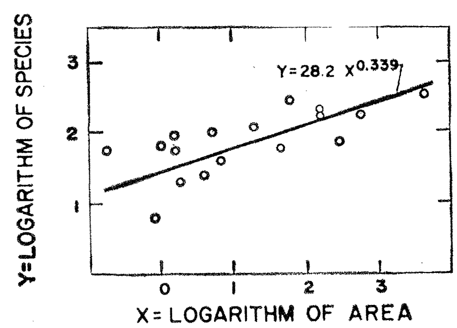

Fig. 2. Relation of species number to area, plotted to logarithmic scales.

is good enough for 1000 species. (These are approximately equal log unit errors.)

Statistics is a powerful tool for sorting out variables, if an underlying model is known. Statistics cannot find the model. The model depends upon the scientist and how he thinks the system reacts. I think model 2 is superior statistically to model 1. Perhaps "groups . . . previously studied by the Arrhenius approach (model 2) need to be examined by multiple regression analysis." However, before any such examination is made, the scientist must (i) state his assumptions, both physical and statistical; (ii) state his linear hypothesis and defend it *before* analysis (both models 1 and 2 are linear in the statistical sense); and (iii) use statistical analysis to test hypotheses, not to find them.

D. R. DAWDY

2350 Harvard Street,
Palo Alto, California

5 June 1964

Dawdy fails to state explicitly that our study was a multifactorial analysis wherein was tested the idea that insular number of land plant species (Y) in the Galapagos Archipelago varies from island to island in such a way that Y depends upon several X's (area, elevation, isolation, and so forth) for its value. We compared multiple regression of Y on 5 X's by model one and model 2, and concluded for multiple—not single—regression that the first model gives better predictions of Y. We noted in our report that area by model 2 gives better predictions than by model 1. This is the single-regression matter to which Dawdy addresses himself. Thus we are delighted to see his confirmation and amplification in the first six paragraphs and figures of his comment of one of our several findings.

Several of his points, nevertheless, require clarification:

1) I am surprised that Dawdy considers our statement of use of the contributions to R^2 of the X's "no exact statement of [our] test of 'goodness of fit'." R^2 or the coefficient of multiple determination is the fraction of total sum of the squared deviations from the mean of Y which is attributable to regression ($\Sigma\hat{y}^2/\Sigma y^2$). As such its value varies from 0 to 1 or perfect prediction. It is widely used in statistics to compare the relative predicting powers or "goodness of fit" for different models or estimating equations. For example, such values for the single-regression, species-area problem here discussed are as follows: By model 1, r^2 (= coefficient of determination) for Y regressed on X is 0.33; by model 2, r^2 for log Y on log X is 0.58.

2) The regression line in Dawdy's Fig. 2 is labeled incorrectly. That "linear" line should be labeled $\log \hat{Y} = 0.1456 + 0.331 \log X$. His designation, $y = 28.2\ x^{0.331}$, represents a transformation to arithmetic values of the preceding equation, and gives in fact a "curvilinear" line fitting the Y points of his Fig. 1. Thus model 1 and model 2 are respectively linear and curvilinear in the "mathematical sense," if not in the "statistical sense."

The fit of predictions ($\hat{Y}$) of Y by model 2, although better than by model 1, is not very good. And to compare the log-log plotting (without conversions of the predicted, log $\hat{Y}$ values to their arithmetic counterparts) with the arithmetic-to-arithmetic one (with 4/5 of the space above the X axis given to one large sample item) could be misleading. My Fig. 1 compares the prediction errors (Y minus $\hat{Y}$) for species numbers regressed on area by both models. That neither is very good can readily be seen. (Close predictions would result in the error points or deviations being restricted to the broken, zero line crossing the figure.) It was at this point in the preliminary work on our problem that we shifted from single to multiple regression. There we found R^2 values of 0.84 and 0.67 for the respective multiple regressions of Y on the 5 X's and of log Y on the 5 log X's. Since we were seeking evidence for environmental control of species abundance by the finding of environmental correlates of insular species numbers [furthermore, G. W. Snedecor (*Statistical Methods*, Iowa State College Press, Ames, ed. 5, 1957, p. 438) suggests not speculating about cause and effect unless R^2 is above 0.80], we concluded in favor of the higher R^2 value (*viz.*, model 1 in multiple regression).

3) Dawdy's stricture that the scientist "must" use statistical analysis to test hypotheses, not to find them, is to me unacceptably typological. I think the majority of actively working scientists would agree that hypotheses are where one finds them. I see in our disagreement the old conflict between induction and deduction. Paraphrasing the writing of a recent commentator on the matter, I suggest that scientists steer a course whose path passes carefully between the twin dangers of accumulating only facts by induction and of resorting to excessive speculation without checks by observation and experimentation.

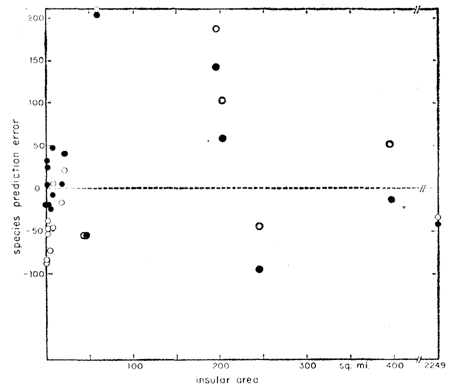

Fig. 1. Prediction errors for species numbers regressed on area. Error = $Y - \hat{Y}$: ○ for $\hat{Y} = 95.44 + 0.117\ X$ where $r^2 = 0.33$; ● for $\hat{Y} = 0.286\ X^{0.331}$ where r^2 for $\log \hat{Y} = 1.456 + 0.331 \log X$ is 0.58.

Finally and since we are involved here in controversy, it is not inappropriate for me to poke fun at myself and Dawdy by quoting John Kenneth Galbraith's view (*Economics and the Art of Controversy*, Vintage, New York, 1959, pp. 104–5) of the nature of controversy in another field of inquiry: "Indeed, the intensity of the debate may be inversely related to the urgency of the questions involved. . . . We should on occasion get a larger perspective on the melee and on the weapons, tactics, and objectives. So viewed, some of the battles will doubtless be shown to be very important. Some, it will be found, are being fought with blank cartridges for ground that has already been won in a war that is over."

T. H. HAMILTON

University of Texas, Austin

26 July 1964