A simplified table for the estimation of magnitude of experimental effect¹

HERBERT FRIEDMAN, College of William and Mary, Williamsburg, Va. 23185

The effectiveness of the independent variable in an experiment may be measured by the correlation between group membership and the dependent variable (r_m). A concise table is presented for the estimation of the value of r_m directly from the probability value for an inferential test and the sample size. The table is well-suited for use in an undergraduate statistics course and can be used with both parametric and nonparametric tests. The table may also be employed to aid in the estimation of appropriate sample sizes.

In a recent paper a table was presented for the estimation of r_m , a measure of the magnitude of experimental effect (Friedman, 1968). This measure of the effectiveness of the independent variable is the correlation between membership in the experimental groups and the dependent variable or score (cf. Cohen, 1967). When the t test is the appropriate measure for the difference between the two groups, r_m is the point-biserial correlation (r_{pb}) which is a function of the value of the statistic, t , and a measure of sample size, df :

$$r_{pb} = \sqrt{\frac{t^2}{t^2 + df}}$$

Analogous formulations are available for χ^2 and for the F test. The magnitude of effect measure may be considered an objective value which relates to "practical significance" in contrast to the p value which relates to statistical significance.

This material has been successively incorporated in an undergraduate introductory statistics course. However, the original table could not readily be used for several important nonparametric statistical tests. To simplify the use of the table, to make it smaller and suitable for all tests, the table was recast in terms of the probability value for the statistic (rather than the statistic itself) and a measure of the sample size. The final version is shown in Table 1.

The appropriate measure of sample size for several common statistical tests is given in Table 2.

Table 1 is based on the t test being used with each sample size and giving the appropriate probability value. It is clear that r_m can also be interpreted in terms of the sample size needed to detect the difference between the groups, i.e., the larger the sample size needed to detect a difference, the smaller the difference.

The level of r_m which would be considered to be large or small is necessarily a function of the particular situation. Where the relevant variables are well controlled, as in a typical Skinner box experiment, and when powerful factors such as drive level are manipulated, relatively large values of r_m would be expected. When many variables contribute to a situation, e.g., academic achievement, the effective manipulation of any one of these variables would produce a relatively smaller effect. Published studies may often be used to provide representative r_m values. Thus a standard operant conditioning experiment which requires 120 Ss to detect the effect of different drive levels provides evidence that, though the effect is significant, it is also of little practical significance ($r_m = .18$). At the other extreme, a similar experiment with sample size = 10 and $p = .10$ yields $r_m = .50$ which is large enough to suggest that the independent variable may well be effective but that the sample size is too small to permit a clear-cut

Table 1
Magnitude of Experimental Effect (r_m) as a Function of the Sample Size Needed to Detect the Effect at a Given 2-Tailed Probability Level

$p =$ Sample Size	.20	.10	.05	.01	.001
1	.95	.99	1.00	1.00	1.00
2	.80	.90	.96	.99	1.00
3	.69	.81	.88	.96	.99
4	.61	.73	.81	.92	.97
5	.55	.67	.75	.87	.95
6	.51	.62	.71	.83	.92
7	.47	.58	.67	.80	.90
8	.44	.55	.63	.76	.87
9	.42	.52	.60	.73	.85
10	.40	.50	.58	.70	.82
11	.38	.48	.55	.68	.80
12	.36	.46	.53	.66	.78
13	.35	.44	.51	.64	.76
14	.34	.43	.50	.62	.74
15	.33	.41	.48	.60	.72
16	.32	.40	.47	.59	.71
17	.31	.39	.46	.58	.69
18	.30	.38	.44	.56	.68
19	.29	.37	.43	.55	.67
20	.28	.36	.42	.54	.65
21	.28	.35	.41	.53	.64
22	.27	.34	.40	.52	.63
23	.27	.34	.40	.51	.62
24	.26	.33	.39	.50	.61
25	.25	.32	.38	.49	.60
26	.25	.32	.37	.48	.59
27	.25	.31	.37	.47	.58
28	.24	.31	.36	.46	.57
29	.24	.30	.36	.46	.56
30	.23	.30	.35	.45	.55
40	.20	.26	.30	.39	.49
60	.17	.21	.25	.32	.41
120	.12	.15	.18	.23	.29

conclusion. The use of r_m helps the student to appreciate that data interpretation involves more than an arbitrary and dichotomous decision to reject or not reject the null hypothesis. The task is both to determine if the observed outcome may be attributed to chance and to consider the size of the effect. Thus, the data of inferential tests is handled in essentially the same fashion as correlational data.

Table 1 can also be used to help estimate the sample size needed to detect an effect at the desired level of significance. The expected or desired r_m value can be found in Table 1 and the corresponding value of the sample size is indicated. Similarly, if data from a pilot study or initial sample is

Table 2
The Measure of Sample Size for Several Statistical Tests

Test	Sample Size
t test	df
F test	$df_{error \text{ term}}$
U test	$(N_1 - 1) + (N_2 - 1)$
Matched Pair Signed Ranks	$N - 1$
χ^2	N = total observations
Kruskal-Wallis	N = total observations
Friedman Analysis of Variance	$N(k)$ = total observations
Binomial test	N = total observations

(Continued on page 195)

A group of 15 control Ss (male undergraduate students) received the same tests, in the same order, and at the same time intervals as the experimentals.

RESULTS

Two-tailed t-tests for independent measures were performed on the pre-post difference scores of the experimental and control Ss. The results revealed a heightened sensitivity to pressure on *all* skin areas, in the experimental relative to the control Ss, at the end of the one-week period of visual deprivation. Statistically significant results were obtained on the finger ($p < .01$), forearm ($p < .002$), neck ($p < .05$), and leg ($p < .05$), but not on the palm. Furthermore, significant aftereffects were still present on postdeprivation Day 2 for the finger ($p < .01$) and on postdeprivation Day 5 for the forearm ($p < .02$) and neck ($p < .05$).

The results on the remaining three measures, i.e., tactual localization and absolute and differential auditory localization, showed no significant differences between the experimental and control Ss. Furthermore, no indication of a trend toward increased accuracy of localization in the experimental Ss was evident on any of the measures.

DISCUSSION

The results of this experiment have indicated that prolonged visual deprivation can produce a significant increase in pressure sensitivity of the finger, forearm, neck, and leg, with the aftereffects persisting for several days after the restoration of normal visual stimulation. A trend toward increased sensitivity of the palm was also observed but the change was not statistically significant. Essentially similar results, derived from other cutaneous measures, have been demonstrated by Zubek et al (1964a) who reported that a week of darkness produced not only a significant increase in heat and pain sensitivity of the forearm but also an increase in tactual acuity of the index finger, palm, and forearm. Furthermore, the effects persisted for a number of days after the termination of the experimental condition. In view of this general increase in cutaneous sensitivity, as indicated by a variety of measures, it is surprising that tactual localization did not show a significant improvement in accuracy or even a trend toward improvement. One possible explanation of this negative finding may be that tactual localization is much more dependent on practice and learning than are the other cutaneous measures which have been employed (Halman & Wright, 1960). A similar explanation may account for the nonsignificant results on both absolute and differential auditory localization. Held (1955), for example, has shown that experience in localizing a sound for a period of 1 h produced a 10-deg change in the accuracy of localization.

Since practice may be the main factor responsible for producing these negative results, it might be fruitful, in future research, to employ a variety of other auditory measures which are only minimally dependent upon practice and learning, e.g., frequency and intensity discrimination. If this were to be done, the results might indicate that prolonged visual deprivation can produce auditory facilitatory effects of a much more general nature than appears to be the case at present (Duda & Zubek, 1965), particularly since Gibby

(1967) has recently reported that visual deprivation of only 3-h duration can produce a significant improvement in both pitch and loudness discrimination as measured by the Seashore Test of Musical Talents.

The demonstration of a significant increase in pressure sensitivity of the finger, forearm, neck, and leg but no change on the palm is puzzling especially since Zubek et al (1964a) have reported a significant improvement in tactual acuity of the palm (two-point threshold) after a week of visual deprivation. It is important to note, however, that a similar differential pattern has been observed by Weinstein et al (1967) after a four-day period of *tactual* occlusion of a circumscribed area of the forearm and palm. Using a von Frey hair technique, Weinstein reported a significant increase in absolute pressure sensitivity of the forearm but no significant change in the sensitivity of the palm. These results, therefore, appear to suggest that a prolonged period of sensory restriction (either visual or tactual) produces an effect on the pressure sensitivity of the palm which is different from that observed on more proximal or distal areas of the arm. The reason for this anomaly must await further research.

In conclusion, these results, indicating an increase in pressure sensitivity after prolonged visual deprivation, are consistent with one of the predictions proposed by Schultz (1965) in his sensoristatic theory of the nervous system, i.e., "when stimulus variation is restricted, central regulation of threshold sensitivities will function to lower sensory thresholds. Thus, the organism becomes increasingly sensitized to stimulation in an attempt to restore the balance [p. 32]." Furthermore, according to this theory, an improvement in tactual and auditory localization might not be expected since these performance measures largely involve learning rather than threshold determinations of sensitivity.

REFERENCES

- DUDA, P., & ZUBEK, J. P. Auditory sensitivity after prolonged visual deprivation. *Psychonomic Science*, 1965, 3, 359-360.
GIBBY, R. G. The effects of visual restriction on visual and auditory discrimination. *Dissertation Abstracts*, 1967, 27(9-B), 3308.
HALMAN, C. R. E., & WRIGHT, G. H. Tactile localization. *Brain*, 1960, 83, 677-700.
HELD, R. Shifts in binaural localization after prolonged exposures to atypical combinations of stimuli. *American Journal of Psychology*, 1955, 69, 526-548.
SCHULTZ, D. P. *Sensory restriction*. New York: Academic Press, 1965.
SEMMES, J., WEINSTEIN, S., GHENT, L., & TEUBER, H. L. *Somatosensory changes after penetrating brain wounds in man*. Cambridge: Harvard University Press, 1960.
WEINSTEIN, S., RICHLIN, M., WEISINGER, M., & FISHER, L. The effects of sensory deprivation on sensory, perceptual, motor, cognitive, and physiological functions. Technical Report CR-727, March 1967, NASA, Washington, D.C.
ZUBEK, J. P., FLYE, J., & AFTANAS, M. Cutaneous sensitivity after prolonged visual deprivation. *Science*, 1964a, 144, 1591-1593.
ZUBEK, J. P., FLYE, J., & WILLOWS, D. Changes in cutaneous sensitivity after prolonged exposure to unpatterned light. *Psychonomic Science*, 1964b, 1, 283-284.

NOTE

1. This research was supported by the Defense Research Board, Canada (No. 9425-08) and by the National Research Council, Canada (APA-290).

(Continued from page 193)

assumed to provide a good estimate of the true magnitude of effect, then we can use Table 1 to indicate the appropriate total sample size. In the above example with $r_m = .50$, the column for $p = .05$ shows that a sample size = 14, or a total of 16 Ss, would be needed to obtain significance at the .05 level. With the original table additional material and procedures are available for estimating sample size.

REFERENCES

- COHEN, J. Some statistical issues in psychological research. In B. B. Wolman (Ed.), *Handbook of clinical psychology*. New York: McGraw-Hill, 1967. Pp. 95-121.
FRIEDMAN, H. Magnitude of experimental effect and a table for its rapid estimation. *Psychological Bulletin*, 1968, 70, 245-251.

NOTE

1. Supported in part by NASA Research Grant NGR 47-006-028.