across days (df = 9/36) shows a clearly significant effect (p < .01).

It is of incidental interest that, in both the experimental condition and the control condition, the between-Ss effect was significant (df = 4/36, p < .01, in each case).

The results of the present study obviously challenge explanation by any current theory. Our own speculation would be to the effect that the Ss were able to maintain an anlytical attitude and to be somewhat objective, in the experimental condition, during the early stages of their service, but that, as time passed and motivation perhaps weakened, the Ss became less critical and their perceptions less realistic. We might then imagine that their tendency to perceive the standard disc as darker and darker during the experimental phase transferred itself in some degree even to the judgments made in the control condition. Clearly, though, such an argument is a weak, ad hoc thing, and there is no obvious explanation in judgmental terms for the results reported here. It should be pointed out, however, that the data would be at least as difficult to explain at the purely retinal level. What seems to be required is additional, clarifying investigation.

REFERENCES

BENUSSI, V. Zur Psychologie des Gestalterfassens. In A. Meinong (Ed.), Untersuchungen zur Gegenstandstheorie und Psychologie. Leipzig: Barth, 1904.

BERMAN, P. W., & LEIBOWITZ, H. W. Some effects of contour on simultaneous brightness contrast. Journal of Experimental Psychology, 1965, 69, 251-256.

DAY, R. H. The effects of repeated trials and prolonged fixation on error in the Müller-Lyer figure. Psychological Monographs, 1962, 76, No. 533.

DEWAR, R. Stimulus determinants of the practice decrement of the Müller-Lyer illusion. Canadian Journal of Psychology, 1967, 21, 504-520.

HURVICH, L. M., & JAMESON, D. Outlines of a theory of the light sense. Cambridge, Mass.: Harvard University Press, 1964. (Translation of: Hering, E. Grundzüge der Lehre vom Lichtsinn. Berlin: Springer, 1920.)

HURVICH, L. M., & JAMESON, D. The perception of brightness and darkness. Boston: Allyn & Bacon, 1966.

JUDD, C. H. Practice and its effects on the perception of illusions. Psychological Review, 1902, 9, 27-39.

KÖHLER, W., & FISHBACK, J. The destruction of the Müller-Lyer illusion in repeated trials: I and II. Journal of Experimental Psychology, 1950, 40, 267-281 and 398-410.

LEWIS, E. O. The effect of practice on the perception of the Müller-Lyer illusion. British Journal of Psychology, 1908, 2, 294-306.

MOUNTJOY, P. T. Effects of exposure time and intertrial interval upon decrement to the Müller-Lyer illusion. Journal of Experimental Psychology, 1958, 56, 97-102.

PARRISH, M., & SMITH, K. Simultaneous brightness contrast as a function of perceptual set. Psychonomic Science, 1967, 7, 155-156.

RATLIFF, F. Mach bands: Quantitative studies on neural networks in the retina. San Francisco: Holden-Day, 1965.

Generalization of a voluntary response as a function of presentation frequency of the training stimulus in testing[1]

JOHN A. HEBERT,[2] Colorado State University, Ft. Collins, Colo. 80521, and JACK E. CAPEHART, University of Arizona, Tucson, Ariz. 85721

Frequency of presentation of the original training stimulus in the test phase of a study of human voluntary generalization was found to affect the form of the generalization gradient on a weight dimension. That is, in a situation in which all of the additional test stimuli are larger than the original training stimulus, and all stimuli are presented with equal frequency, a tendency to choose stimuli toward the middle of the test range as the original was observed. However, if the original stimulus was presented more frequently than any of the other stimuli in testing, Ss tended to choose the original "correctly," yielding a "typical" unidirectional generalization gradient.

In a voluntary response generalization test, it is not unusual to present the original training stimulus (S1) more frequently in the test phase than the additional test stimuli. This procedure of "overrepresenting" S1 in testing, or the procedure of "booster" trials with S1 interspersed with test trials, may actually produce generalization gradients quite different from those obtained with a procedure in which S1 is presented with the same frequency as each of the test stimuli, i.e., S1 is treated as just another test stimulus. Examination of the literature points to such a possibility. A study by White (1965) illustrates the overrepresentation procedure. White trained human Ss to respond to a red-colored stimulus by pushing a lever upward. Ss were also instructed to push the lever downward if they saw any stimulus other than the red one (S1). A test series consisted of 16 presentations of S1 interspersed with three presentations each of the additional test stimuli (S2 through S5). All of the additional test stimuli were in the direction of shorter wavelengths from S1 (i.e., unidirectional testing). White found that response was greatest to S1 and that it decreased as similarity to S1 decreased, i.e., a "typical" unidirectional gradient. However, a study by Thomas & Jones (1962), illustrating the equal presentation procedure, produced somewhat different results. Ss were exposed to a 525-mu stimulus for a brief time and were told to release a key only if they saw that value again in a test series. In the groups receiving unidirectional testing, that is, test stimuli being either all shorter or all longer than S1 in wavelength, Thomas and Jones reported that Ss tended to choose stimuli in the center of the test range (e.g., S2 or S3) more than the decentered S1. Instead of a gradient with a peak at S1 with a steady decrement to S5, as in the White study, Thomas and Jones reported a peak at S2 with a decrement in either direction. While these two studies are really not very comparable because of many minor differences, it is possible that the differences in presentation frequency of S1 in test is the reason for the difference in the gradient form. A replication of the Thomas and Jones' finding of a "central-tendency" by Helson & Avant (1967), with a size dimension and an equal presentation procedure, makes this hypothesis more compelling.

This present study investigates the effect of "overrepresentation" of S1 in testing. A wavelength dimension was not chosen because of the possibility of S using verbal labels in the task of identifying stimuli. The label "red," for example, is an appropriate label for wavelengths of specific value and, conceivably, could help S to tie down the stimuli. Work by Thomas and his coinvestigators (Thomas & Mitchell, 1962; Thomas & Bistey, 1964; Thomas &
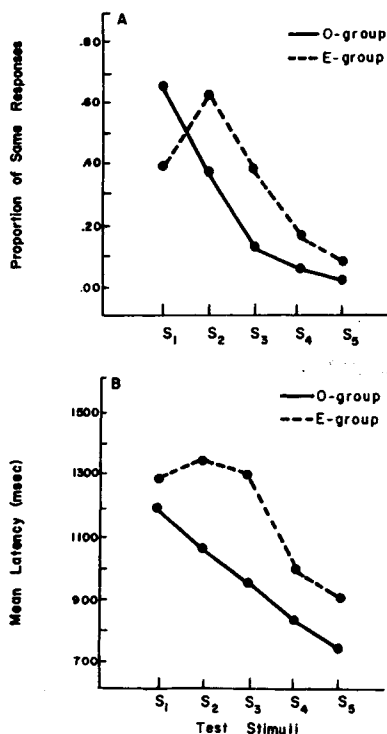
Fig. 1. (a) Proportion of "same" responses to the test stimuli for the combined O- and E-groups. (b) Decision time to the test stimuli for O- and E-groups. (S1 = 100 g, S2 = 110 or 125 g, etc.)

DeCapito, 1966) indicates clearly that Ss' labeling responses lead to predictable shifts in response tendencies. The dimension of lifted weights seemed appropriate for this problem since a verbal label such as "light" is not tied to a specific value of the dimension.

METHOD

Forty male and female students of introductory psychology courses at the University of Arizona served as Ss. Two ranges of five weights each were used as stimuli. One distribution ranged from 100 to 140 g in 10-g steps, while the other ranged from 100 to 200 g in 25-g steps. Stimuli were made from aluminum cans 4 in. tall and 2 in. in diam, and were all identical in appearance. When a single weight was lifted with the preferred hand of S, it released a microswitch which started an electronic counter-timer. When S reached his decision, he pressed one of two response keys with the free hand, which stopped the counter-timer and recorded the choice and latency of response by means of a digital recorder. The response was made with the index finger of the free hand which rested between, and slightly above, the two closely spaced response keys.

Subjects were trained on the same stimulus (100 g) and then given one of four test series: Group E-10 was tested with the 10-g series and was presented with S1 with

equal frequency as the other stimuli (S2 through S5); Group E-25 was tested with the 25-g series and received equal presentation of S1; Group O-10 was tested with the 10-g series and was presented with S1 five times more often than any of the other stimuli; Group O-25 was tested on the 25-g series and received S1 five times more frequently than the others. The equal presentation (E) groups were given five blocks of trials in which each of the five stimuli were presented once per block. The overrepresentation (O) groups were also given five blocks of trials, but within each block S1 was presented five times while each of the other stimuli were presented only once. A three-way repeated-measures design was used, with testing procedure (O vs E), stimulus interval (10 vs 25 g), and stimuli as the main variables.

Upon entering the experimental room, Ss were informed that they were participating in a weight-perception experiment. They were instructed to lift S1 (100 g) five times and told that they were to identify it in a test series of weights to be presented one at a time. They were instructed to push a "same" key if they felt it was the same as S1, and to push a "different" key if they felt it was not. Both speed and accuracy were urged.

RESULTS

The hypothesis that procedural differences can produce differences in the gradient shape is confirmed. Figure 1 depicts the pattern of both choice and latency data for the combined E groups and combined 0 groups. Since the effect of stimulus interval (10 vs 25 g) was not significant for both dependent measures (F = 1.98, df = 1/36, p > .05 for choice; F = .63, df = 1/36, p > .05 for latency), the data in Fig. 1 are combined across the stimulus interval variable. The choice data expressed in the proportion of "same" responses (Fig. 1a) show that the 0 groups, represented by the solid line, chose S1 most frequently and made fewer "same" responses as the test stimuli became more dissimilar to S1, i.e., a gradient similar to those of White. The E groups (dotted line), on the other hand, chose S2 most frequently, and "same" responses decreased in both directions as a function of similarity to S2, i.e., similar to the gradients of Thomas and Jones. The Groups by Stimuli interaction shown in Fig. 1a is significant (F = 5.74, df = 4/144, p < .01). It may be argued, however, that the significant interaction is due to a steepening of the gradient in the 0 groups as a result of more experience with S1. Trend analysis of the individual groups makes this interpretation less plausible. All four groups produced highly significant linear trends (O-10: F = 23.20, df = 1/36; O-25: F = 31.67, df = 1/36; E-10: F = 13.24, df = 1/36; E-25:

F = 18.87, df = 1/36; all p < .001), but only the E groups produced significant quadratic trends as well (E-10: F = 4.18, df = 1/36, p < .05; E-25: F = 7.46, df = 1/36, p < .01). No other trends were significant. The quadratic trends in the E groups, of course, support the hypothesis that presentation frequency of S1 affects the shape of the generalization gradient.

The latency data (Fig. 1b) closely parallel the choice data; longest latencies tend to be associated with the most often chosen stimuli. It must be noted that these points represent all decisions (both "same" and "different" responses), since reliable curves for "same" responses were hard to obtain because of relatively few responses to the distant stimuli. The interaction of Groups by Stimuli in Fig. 1b is also significant (F = 2.84, df = 4/144, p < .05). The theoretical significance of the correlation of the latency and choice responses is not apparent at this time.

The role of context (asymmetrical test range) should be acknowledged here. Thomas & Jones (1962) and Helson & Avant (1967) did not observe the central-tendency effect when S1 was placed in the middle of the test range (i.e., bidirectional testing). Therefore, one might conclude that frequency effects in this study interacted with context effects to produce the results. A design which can assess the effects of frequency independently of context is needed.

REFERENCES

HELSON, H., & AVANT, L. L. Stimulus generalization as a function of contextual stimuli. Journal of Experimental Psychology, 1967, 73, 565-567.

THOMAS, D. R., & BISTEY, G. Stimulus generalization as a function of the number and range of generalization test stimuli. Journal of Experimental Psychology, 1964, 68, 599-602.

THOMAS, D. R., & DeCAPITO, A. Role of stimulus labeling in stimulus generalization. Journal of Experimental Psychology, 1966, 71, 913-915.

THOMAS, D. R., & JONES, C. G. Stimulus generalization as a function of the frame of reference. Journal of Experimental Psychology, 1962, 64, 77-80.

THOMAS, D. R., & MITCHELL, K. Instructions and stimulus categorizing in a measure of stimulus generalization. Journal of the Experimental Analysis of Behavior, 1962, 5, 375-381.

WHITE, S. H. Training and timing in the generalization of a voluntary response. Journal of Experimental Psychology, 1965, 69, 269-275.