

Voluntary stimulus generalization as a three-category judgment process*

FRANK D. McGUIRK and JOHN A. HEBERT
Colorado State University, Fort Collins, Colo. 80521

Human Ss were asked to heft a 150-g weight and were required to identify it in a test series of weights, ranging in value from 100 to 200 g, with 25-g steps. Following each test response, each S was asked to rate the weight he had just lifted on a 7-point scale of heaviness, with 1 being very light, 4 neither light nor heavy, and 7 being very heavy. Results showed that Ss chose the stimulus weight closest to adaptation level (125 g) as the original training stimulus more often than the 150-g weight itself. Support was noted for adaptation-level theory. Further, analysis of rating data yielded information as to why Ss rejected a given stimulus as not the original. Three gradients of response, "too light," "same," and "too heavy," were drawn, showing this voluntary generalization task to be a three-category judgment process.

In a study of voluntary stimulus generalization, S is usually trained on a single stimulus value, S+, and then required to identify S+ in a test series of stimuli along the same dimension. Normally S is required to respond according to one of two basic methods. In the single-response method, S is to make an active response, R₁, if he thinks a stimulus is the same as S+, and to *inhibit* R₁ if he thinks it is different from S+. With the single-response method, for example, S might be required to release a key if he thinks S+ is present but not release it if he thinks a different stimulus is present (e.g., Thomas & Jones, 1962). A second approach, the two-response method, is similar to the single-response method except that, in addition to requiring S to give R₁ to stimuli he equates with S+, a different and separate response, R₂, is required of him if he thinks the stimulus is different from S+. The two-response method, for example, might require S to press a "same" key if he thinks S+ is present and a "different" key if he thinks the stimulus is not S+ (e.g., Hebert & Capehart, 1969).

The two-response method yields more information about the judgment process than does the single-response method, since with the former method one can measure how long it takes a given S to decide that a stimulus is different from S+. On the other hand, with the single-response method E uses a fixed criterion of elapsed time (e.g., 2 sec) to define the inhibition of R₁, but really has no idea *when* the decision to inhibit the response was made.

Neither of these two methods yields complete information about the judgment process, however, because E does not find out *why* S rejects a given stimulus as not S+. For example, if S+ is, say, 150 g, does S reject a 125-g test stimulus as too light or too heavy? Presumably, 125 g would be rejected as too light on most occasions, but it is possible that S may occasionally reject this stimulus as too heavy. Obtaining this information from S would afford a more complete picture of the judgment process, since E could plot *three* gradients of response: a gradient of "too light" responses, a gradient of "same" responses (the generalization gradient), and a gradient of "too heavy" responses. The present study of voluntary stimulus generalization along a weight dimension was conducted in such a manner, thus affording this more complete analysis of the judgment process.

METHOD

Nine introductory psychology students served as Ss in order to fulfill a course requirement. The stimulus weights, ranging in value from 100 to 200 g in 25-g steps, were aluminum cans 4 in. tall and 2 in. in diam which were filled with wax and shot and were painted black to be identical in appearance. When a single weight was lifted with the preferred hand of S, a microswitch was released which started a Standard timer. When S reached his decision, he pressed one of two response keys ("same" or "different") which stopped the timer. Both choice and latency of response were then recorded by E. The response was made with the index finger of the free hand, which rested between and slightly above the two closely spaced response keys. Since S lifted all weights in a box-like housing, all visual cues were eliminated.

Upon entering the experimental room, Ss were informed that they were participating in an experiment involving weight perception. After familiarization with the experimental situation, Ss were instructed to heft the 150-g weight (S+) five times and were told to identify it in a subsequent test series of weights to be presented one at a time. They were instructed to push a "same" key if they thought a weight was the same as S+ and a "different" key if they felt it was not S+. All Ss were urged to make this response as accurately and quickly as possible. After each response, Ss were asked to rate the stimulus they had just lifted on a 7-point scale (1 = very light, 2 = light, 3 = fairly light, 4 = neither light nor heavy, 5 = fairly heavy, 6 = heavy, and 7 = very heavy) and to call out the rating number appropriate to their judgment. For S's convenience, the verbal and numerical equivalents of this scale were displayed at eye level on the housing. All Ss were given five blocks of test trials in which each of the five test stimuli (100, 125, 150, 175, and 200 g) were presented once per block, for a total of 25 test trials. The order of presentation within each block was systematically varied between Ss, and Ss were assigned at random to the orders.

RESULTS AND DISCUSSION

Figure 1 depicts the patterns of choice and latency of

*Requests for reprints should be sent to John A. Hebert, Department of Psychology, Colorado State University, Fort Collins, Colorado 80521. The authors thank Henry Cross for his reading of the manuscript.

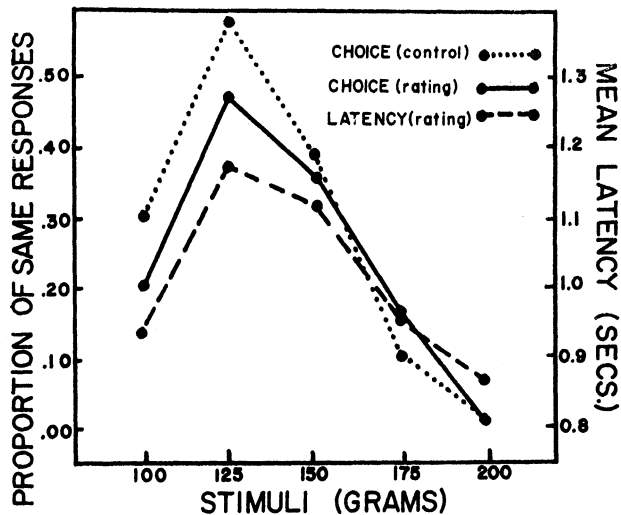


Fig. 1. Gradients of latency and choice of response to the test stimuli.

Table 1
The Proportion of "Same" Responses and the Mean Category Ratings of the Test Stimuli

Stimulus Weights (Grams)	Mean Category Rating	Proportion of "Same" Responses
100	2.42	.20
125	3.87	.47
150	4.51	.38
175	5.40	.16
200	6.20	.02

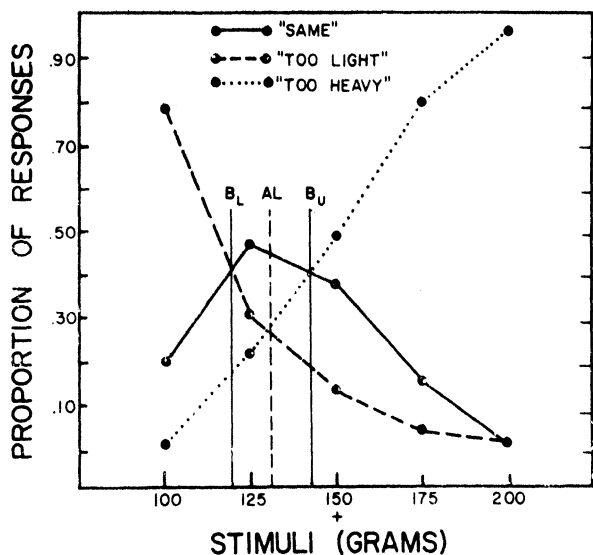


Fig. 2. Gradients of the proportion of "too light," "same," and "too heavy" responses to the test stimuli, with category boundaries and AL drawn in.

(generalization) responses, not the 150-g stimulus as might be expected. This finding is in agreement with the adaptation-level (AL) theory proposed by Capehart, Tempone, & Hebert (1969). An AL explanation of these data goes as follows. In training on a single stimulus value, 150 g, an AL is established at that value represented on S's psychological continuum. What S learns is not "stimulus of 150 grams is correct" but "stimulus value which corresponds to AL is correct." The introduction of these particular test stimulus values (100-200 g in 25-g steps) creates a slight shift in AL. (The AL of the test series is calculated to be approximately 128 g, according to Helson's (1947) formula for lifted weights.) According to the theory, 125 g will be chosen most often as S+, since it is closest to the test AL. Mean heaviness ratings obtained from Ss after the "same"- "different" response further support the assertion that 125 g is closest to AL. Using a linear interpolation method, the stimulus which would have received a neutral "4" rating was 131 g, a value which can be taken as a direct estimate of AL (Helson, 1947). Clearly, 125 g is closer to the estimated AL than any other test stimulus. Table 1 gives the mean category rating for each stimulus value, as well as the proportion of "same" responses to each stimulus.

A question arises, however, about the task of rating the stimulus after making the "same"- "different" judgment and its effect on the subsequent generalization responses. The control gradient (dotted line) in Fig. 1 depicts the generalization gradient from an unpublished study¹ which is highly similar to the present study but which did not require ratings from Ss. The pattern of responses is clearly the same, although a higher level of responding is present in the control gradient. It cannot be determined if rating has any impact on the overall level of response in the present study.

A more refined analysis of the data is necessary to find out why S rejected a given stimulus as not S+. Since 50 of the 54 "same" responses were subsequently rated "4" (neither light nor heavy) and since 50 of the 51 "4" ratings were associated with "same" responses, it is clear that Ss associated the correct stimulus with AL or the neutral rating, even though no instructions to do so were given by E. This finding supports the basic assumption of the AL theory of Capehart et al (1969), namely, that S learns that the "stimulus value which corresponds to AL is correct," and also allows for an accurate analysis of the "different" responses. For example, if S rated the 125-g stimulus "5," "6," or "7" after pressing the "different" key, one can assume he rejected it as being too heavy. By the same rationale, if he rated the 125-g stimulus "3," "2," or "1" after calling it "different," he can be said to have rejected it as too light. Using this analysis, the gradients in Fig. 2 were obtained. Note the points at which the gradients cross in Fig. 2. The crossing of the "too light" gradient (dashed line) with the "same" gradient (solid line), as well as the crossing of the "same" gradient with the "too heavy" gradient (dotted line), could be taken as indicators of the upper

response in this experiment. Inspection of the choice gradient (solid line) reveals that the 125-g stimulus elicited the greatest proportion of "same"

and lower boundaries of the "same" category (B_L and B_U , respectively, in Fig. 2). What emerges from this graphical analysis is a clear picture of stimulus generalization as a three-category judgment process: Does Stimulus X belong in the "too light," "same," or "too heavy" category? Dropping lines perpendicular to the abscissa from the crossing points of the three gradients (as in Fig. 2), one can locate values for B_L and B_U on the physical dimension. Values thus obtained ($B_L = 120$ g, $B_U = 142$ g) are drawn in in Fig. 2, along with the estimated $AL = 131$ g. The fact that the category boundaries, B_L and B_U , are equidistant from AL is noteworthy. This suggests that the midpoint of the "same" category may be defined by AL .

Analysis of the latency data (dashed-line gradient in Fig. 1) can be made in terms of this three-category model. McGuirk (1971) has suggested a simple explanation for the latency data in situations where stimuli are to be categorized: Stimuli which fall close to category boundaries should be associated with long latencies, while stimuli far from category boundaries should be associated with short latencies. In a three-category judgment situation, then, an M-shaped function is predicted, with peaks at each of the two boundaries. Inspection of the latency pattern in Fig. 1 shows no such pattern. However, it should be pointed out that no test values were presented at the boundary values (120 and 142 g), so one could hardly observe peaks in latency at those values. The fact that the test

values closest to the boundaries (125 and 150 g) are also associated with the longest latencies in this study could be interpreted as consistent with McGuirk's (1971) hypothesis. A procedure in which the critical region is sampled more adequately, e.g., an increase in the number of test stimuli, is needed before an understanding of the latency patterns can be reached. A study which not only adequately samples the stimulus dimension but also assesses directly the effect of rating on the generalization gradient is needed. Additionally, a design in which E simply asks S why a stimulus is not the "same" should also be done.

REFERENCES

- Capehart, J., Tempone, V. J., & Hebert, J. A theory of stimulus equivalence. *Psychological Review*, 1969, 76, 405-418.
- Hebert, J. A., & Capehart, J. E. Generalization of a voluntary response as a function of presentation frequency of the training stimulus in testing. *Psychonomic Science*, 1969, 16, 315-316.
- Helson, H. Adaptation level as a frame of reference for prediction of psychophysical data. *American Journal of Psychology*, 1947, 60, 1-29.
- McGuirk, F. D. Latency patterns in category judgments. Unpublished MS thesis, Colorado State University, 1971.
- Thomas, D. R., & Jones, C. G. Stimulus generalization as a function of the frame of reference. *Journal of Experimental Psychology*, 1962, 64, 77-80.

NOTE

1. Hebert, J. A., & Groves, K. Voluntary stimulus generalization as a function of presentation frequencies of stimuli in testing. Unpublished study, 1969.

(Received for publication September 25, 1972.)