Association time, commonality, and the test-retest reliability of free association responses [1]

S. JOYCE BROTSKY[2] AND DONALD C. BUTLER, *SAN FERNANDO VALLEY STATE COLLEGE*

MARIGOLD L. LINTON, *SAN DIEGO STATE COLLEGE*

One hundred and eight Ss gave verbal associates to 68 words presented via a memory drum. Association time (AT) was recorded and the associates were scaled for commonality (C). A week later, 40 of the Ss were presented the same verbal stimuli to determine the probability of occurrence of an identical associate (IA). The results supported two versions of Marbe's Law: (a) Words to which common responses are given are responded to quickly (a within-S relationship). (b) Ss who respond with common associates respond quickly (an across-S relationship). Strong correlations were found between C and IA, both within and across Ss, and between AT and IA within Ss; a weaker correlation was found between IA and AT across Ss.

Marbe's Law has seemed so much a part of the fabric of experimental psychology that further studies of the "law" may seem gratuitous. Investigation of this "law," however, suggests that it has hardly passed beyond the status of a hypothesis, since there have been few attempts to verify it or test its limits. The only adequate study, according to contemporary standards of research, was conducted by Schlosberg & Heineman (1950) who found a correlation of -.80 between commonality (C) and log association time (log AT). However, in their study different Ss contributed AT scores to different stimulus words; hence, a correlation of AT and C calculated across words was partly confounded by subject differences. The present study distinguishes two possible versions of Marbe's Law: (1) Words to which common responses are given are responded to more rapidly (a within-S relationship), and (2) Ss who give common responses respond more rapidly (an across-S relationship).

A related, though separate, problem is that of test-retest reliability. When any frequency based index such as C or frequency of occurrence of the most common response (dominance of the primary) is used to predict individual behavior, the assumption is made that there is a lawful relationship between the response hierarchy derived from a normative group and the hierarchy of the individual S. If this assumption is warranted, it would be predicted that high C, short AT responses would tend to be repeated by a given S on two or more separate occasions. Following the work of Hall (1966), Brotsky & Linton (1967) investigated the probability of Ss' giving identical associations (IA) to 148 words on two separate occasions. Although the test-retest reliability was, in general, much lower than that reported by Hall (.32 vs .50), a correlation of .88 was found between dominance of the primary and the number of IAs. The present study extended this work in exploring the relationships among C, AT, and IA.

Subjects

The Ss were 54 male and 54 female undergraduates between the ages of 17 and 23. All Ss had been born in continental U.S.A. and had learned English as their first language. The Ss participated as part of a course requirement at San Fernando Valley State College.

Stimulus Material

The stimulus materials consisted of two randomized orders (List 1 and List 2) of 68 words chosen from the Connecticut Free Associational Norms (Bousfield, Cohen, Whitmarsh, & Kincaid, 1961). The words were typed in upper case letters on vellum paper and presented to S via a Stowe memory drum.

Experimental Design

Group A consisted of 10 females and 10 males run by E_1; Group B consisted of 24 females and 24 males run by E_2; and Group C consisted of 20 males and 20 females run by E_2. Half of the Ss in each of these six subgroups responded to List 1 while the other half responded to List 2. Group C differed from the other two groups in that each of the Ss was run on two separate occasions, separated by seven days. Those Ss who received List 1 in Phase 1 received List 2 in Phase 2; those Ss who received List 2 in Phase 1 received List 1 in Phase 2. The same procedure was followed in Phase 1 and Phase 2 of the experiment.

Procedure

The stimulus words were presented serially to S at the rate of one every 6 sec. The Ss were given standard free association instructions which emphasized the necessity of responding as quickly as possible with the first word that came to mind. An electric timer was activated by the rotation of the drum as each stimulus word appeared, and stopped when E tapped a telegraph key at the onset of S's response. The E recorded S's verbal associate and his AT to the nearest .01 sec.

Results

A list of associates and their frequencies was developed from the responses of the 108 Ss to the 68 stimulus words during Phase 1. Each associate was then rated for C, defined as the number of times that associated occurred to any given stimulus, divided by the total number of responses to that stimulus. The C scores ranged from .01 for idiosyncratic responses to .91 for THREAD, the primary response to SPOOL. C was also calculated for the responses in Phase 2, but was based on the frequency tallies of Phase 1. Since there were a number of associates in Phase 2 which had not appeared in the Phase 1 response hierarchy for the stimulus word in question, there were some associates that were given a C score of 0 in Phase 2.

Both AT and C were transformed to logarithms.

This transformation is standard for reaction time measures, and was supported for C by the observation of positive skewness of within-cell distributions and a positive correlation of the means and standard deviations of the six subgroups ($r = .57$). All further processing was based on log AT and log $(C + 1)$. No difference was found in log AT as determined by two Es ($t = 1.50$, $df = 100$, $p > .05$); accordingly, the data from E_1 and E_2 were combined.

All the data from Phase 1, based on 108 Ss and 68 stimulus words, were employed in the determination of the correlation between log AT and log $(C + 1)$. This can be computed in two quite different ways: (1) For every stimulus word determine the average AT and C across Ss, and then correlate these two values across stimulus words. This correlation measures the tendency for words to which common responses are given to be responded to more rapidly, and is probably the relationship originally intended by Marbe's Law. Determined this way, $r = -.84$ ($df = 66$, $p < .01$). (2) For every S determine the average AT and C across stimulus words, and then correlate these two values across Ss. This correlation measures the tendency for individuals who, on the average, have long reaction times, to give common responses. Determined thus, $r = -.36$ ($df = 106$, $p < .01$). Clearly, the within-S version of Marbe's Law is a more powerful predictor than the across-S version.

The data from Phase 1 were combined and the mean log AT was found to be 1.96 sec. An analysis of variance performed on log AT to determine the effects of List and Sex indicated that neither the main effects of List, Sex, nor their interaction was statistically significant. A similar analysis of log $(C + 1)$ revealed no significant effects. The error variance for each of the two analyses was based on $df = 104$.

The data from Group C, which consisted of 20 males and 20 females who participated in both Phase 1 and Phase 2, were analyzed further. Analyses of variance were performed on log AT and log $(C + 1)$ to determine the effect of Words, Sex, Phase, and their interactions. A total of 5440 observations were involved, with $df = 36$ for the error variance between Ss and $df = 4860$ for the error variance within Ss. For log AT, significant differences were found among Words ($F = 14.58$, $df = 67/4860$, $p < .01$). As expected, the average AT was longer in Phase 1 (1.94 sec) than in Phase 2 (1.69 sec). This difference was highly reliable ($F = 343.86$, $df = 1/4860$, $p < .01$). Also significant were the interactions between Sex and Words and between Phase and Words. These interactions have F values of 1.53 and 1.32, respectively, and although technically significant are difficult to interpret experimentally. For log $(C + 1)$ the picture was simpler. Only Words ($F = 45.15$, $df = 67/4860$, $p < .01$) and the interaction between Sex and Words ($F = 1.65$, $df = 67/4860$, $p < .01$) were significant. This interaction reflects a complex pattern of differences between the sexes

in the average commonality of their associations to different words. The mean commonality of response was 14.8 in both Phase 1 and Phase 2.

For the 40 Ss who participated in both Phases 1 and 2 of the experiment, the occurrence of IAs in Phase 1 and Phase 2 was recorded. Agreement between Phase 1 and Phase 2 was 60% for male Ss and 64% for female Ss. These values were substantially higher than those found by the authors in an earlier study (Brotsky & Linton, 1967) but more in line with the results of Hall (1966), who found about 50% IAs on test and retest. The high percentage of test-retest agreement in the present study may be due to some combination of long list, short test-retest period, and individual-oral administration procedure.

From the array of 68 words by 40 Ss, it is possible to calculate correlations between log AT and IA, and between log $(C + 1)$ and IA, either across words or across Ss. The correlation between S's average log AT and his total IA was $-.33$ ($df = 38$, $p < .05$), while the correlation between S's average log $(C + 1)$ and IA was $.79$ ($df = 38$, $p < .01$). The results indicated that the commonality characteristic of a particular S's associates was an excellent predictor of whether that S would tend to give many IAs on the retest. The correlation across words of the average log AT and the total IA for each word was $-.79$ ($df = 66$, $p < .01$), while the correlation between log $(C + 1)$ and total IA was $.89$ ($df = 66$, $p < .01$). Thus, both the mean commonality of response and the mean AT to a word provided excellent prediction of whether that word would produce the same associate in Phase 2 as in Phase 1.

Predictions of verbal behavior can be made on the basis of either the characteristics of the Ss or the characteristics of the verbal stimuli. In older studies of C and AT, the confounding of S and stimulus variables has prevented evaluation of relative strength of such differential predictions. The results of the present study, where the effects of these variables were isolated, strongly suggest that superior predictions may be made on the basis of stimulus variables.

References

BOUSFIELD, W. A., COHEN, B. H., WHITMARSH, C. A., & KINCAID, W. D. The Connecticut free associational norms. Tech. Rep. No. 35, Contract N onr 631(00) Office of Naval Research, 1961.

BROTSKY, S. JOYCE, & LINTON, MARIGOLD, L. The test-retest reliability of free association norms. *Psychon. Sci.*, 1967, 8, 425-426.

HALL, J. F. The reliability of free association responses. *Psychon. Sci.*, 1966, 4, 79-80.

SCHLOSBERG, H., & HEINEMAN, C. The relationship between two measures of response strength. *J. exp. Psychol.*, 1950, 40, 235-247.