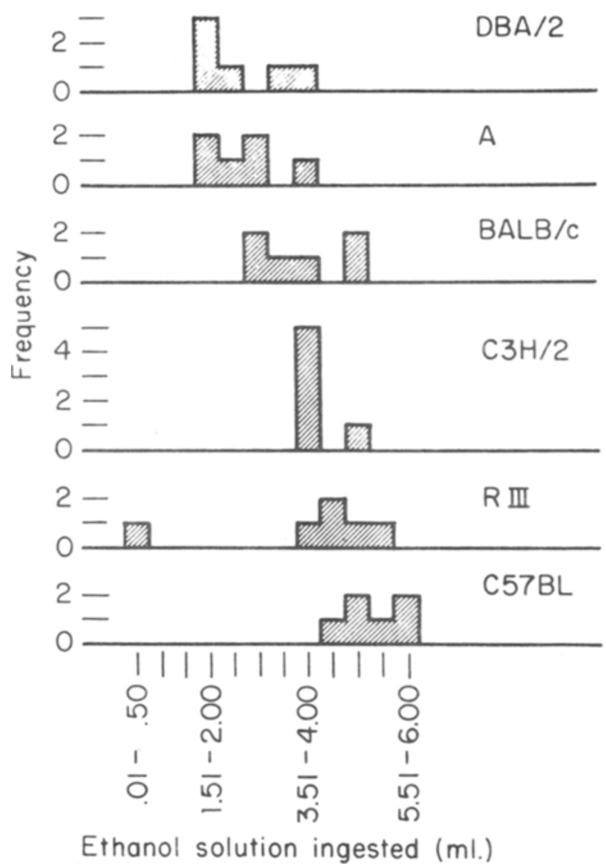
individual cages, and provided with a single cylinder containing tap water. Daily readings were taken of water consumption until Day 3, when the cylinder was removed. Twenty-four hours later, animals were given cylinders containing 10% ethanol; 24 h later, the amount consumed from the ethanol solution under the conditions of thirst motivation was recorded. Figure 3 shows the results. It is apparent that the strain rank ordering obtained by this method is essentially the same as that obtained in the longer 14-day standard test, or in the short version described above.

These results show that some procedures are essentially interchangeable for purposes of assessing alcohol preference, and that another procedure is not. Apart from the practical implications for research on alcohol consumption in animals, these results indicate the usefulness of strain differences in assessing the comparability of different procedures.



REFERENCES

McCLEARN, G. E. Genes, generality, and behavioral research. In J. Hirsch (Ed.), *Behavior-genetic analysis.* New York: McGraw-Hill, Inc., 1967. Pp. 307-321.

McCLEARN, G. E., & RODGERS, D. A. Differences in alcohol preference among inbred strains of mice. Quarterly Journal of Studies on Alcohol, 1959, 20, 691-695.

McCLEARN, G. E., & RODGERS, D. A. Genetic factors in alcohol preference of laboratory mice. Journal of Comparative and Physiological Psychology, 1961, 54, 116-119.

RODGERS, D. A. Factors underlying differences in alcohol preference among inbred strains of mice. Psychosomatic Medicine, 1966, 28, 498-513.

NOTE

1. This research was conducted at the University of California, Berkeley, with the support of National Science Foundation Grant G-14221.

On assessing verbal stimulus hierarchies[1]

CECIL A. ROGERS, JR., *DEPARTMENT OF PSYCHOLOGY, UNIVERSITY OF ARIZONA, Tucson, Arizona 85721*

Typical word-association procedures result in verbal response hierarchies for a specific stimulus word. The problem of assessing corresponding stimulus hierarchies is discussed. Three approaches are entertained: obtaining associations to all of the words in the dictionary, clerical cross-indexing of existing response-hierarchy data, and obtaining backward associations which are then used in forward association. A comparison of data obtainable by the last two methods is presented.

Deese (1962) states that the associative meaning of a word may be defined in terms of (a) the distribution of responses (response hierarchy) to that word when it is used as a stimulus in free association; and (b) the associative distribution (stimulus hierarchy) or collection of stimuli to which the word occurs as a response. Applications of word-association norms for response hierarchies were stimulated by publication of the Russell-Jenkins (1954) norms. Further impetus was provided by the shift in emphasis of interference theories of forgetting from retroactive to proactive agents, as signaled by Underwood (1957) and Underwood & Postman (1960). Norms on verbal-association response distributions have been well-exploited, including studies on paired-associate learning (Coleman, 1963), mediation (Jenkins, 1963), controlled recall (Bilodeau, Fox, & Blick, 1963), sentence recall (Rosenberg, 1966), interference generation (Blick, 1965), semantic generalization (Mink, 1957), and organizational characteristics of free recall (Cofer, 1965).

On the other hand, norms on stimulus distributions have not been generally available for several reasons. First, there is a methodological problem as to how one might assess the stimulus distribution for a given word. While, in assessing response distributions, it is simple enough to ask S to say or write the first word that a given item makes him think of, it is quite another problem to get him to produce a stimulus for some word. Next, even if we could contrive to get Ss to produce a list of stimulus words, the obtained stimulus distribution for a word might not differ from its response distribution. It is well known that many verbal associative networks are tightly knit (e.g., *cold* often elicits *hot* as its primary and vice-versa). Third, while almost every word, when used as a stimulus, produces some response distribution which is fairly reliable and accurate, we might be hard-pressed to think of a procedure which would ensure us that we have sampled all or even many high frequency producers of a given response. Fourth, if somehow we do collect items which Ss say are stimuli for a word, we cannot be sure unless we require them or others to free associate to these stimuli to determine the actual probability with which the response word will occur.

One mode of attack, as suggested by Deese (1962), would be to have large groups of people free associate to all of the words in the dictionary. A gigantic mass of data would result; and, with the aid of a computer, rather complete information would be obtained on response hierarchies for all words. If one is then interested in a stimulus distribution for a particular word, such as *horse*, he may conduct a clerical search to find those response hierarchies in which *horse* is included and the frequency of its occurrence. One could then note whether the stimulus hierarchy for a given word so closely approximates its response distribution as to be of limited additional value. While our technology is adequate for such an undertaking, the enormity of the task makes its feasibility doubtful at present. For example, excluding many slang, dialectic, and archaic words, *Webster's Seventh New Collegiate Dictionary* (1967) still contains over 130,000 entries. If we desired 100 responses for each stimulus word in discrete free association, there would be a need for 13 million pieces of data. Even if each S associated to many stimuli, the required N would be very large.

A second approach, and the only one currently in vogue, is the use of response-distribution norms already on hand. By cross-indexing, one may locate responses which occurred to more than one stimulus word. For example, the Bilodeau-Howell (1965) norms provide response distributions for 180 common words. *Woman* is found to be the primary response to *man* with a probability (p) of .80. A systematic search reveals other stimuli which evoke *woman: lady* (p = .40), *girl* (.07), *pretty* (.03), *queen* (.03), *child* (.02). But there are several problems associated with this procedure. One inadequacy is that very little or no information at all is available for some words. If one searches the Bilodeau-Howell data for elicitors of *chicken*, only two weak stimuli are found: *yellow* (.04), and *brave* (.02). Obviously, the degree of success with this approach is a function of the choice of original stimulus items in the set. Rarely would selection have been made with the purpose in mind of later re-sorting response-distribution data in order to sample from stimulus hierarchies. This clerical re-sorting of data will, then, be promising only if the stimulus words originally chosen for administration were related, and if the word in which one happens to be interested were included.

A third approach, as suggested by Rogers (1965), would be to use a backward-association procedure. *Horse* is presented to a group of Ss with instructions such as: "To the left of the item, write a word which would remind you (or make you think) of that item." Such a procedure is based upon the assumption that Ss can, in recalling pre-experimentally established associations, remember stimuli as well as responses. Our S then searches his repertoire of language habits and selects a stimulus for the item *horse*, such as *Palomino*. The question then arises as to whether or not the so-called stimuli produced by this procedure are really stimuli. Many Ss may free associate rather than backward associate. To find out, one must take these words and re-administer them to determine whether they will indeed evoke the desired response, and if they do, how strongly. For a backward-association technique to work well, there should be a high correlation between frequency with which a stimulus is produced and frequency with which it elicits the response. Should a high correlation obtain, the forward-association step could then be dropped. Failing that, the backward-association procedure should at least produce more effective stimuli than could be found by a clerical sort of existing norms on response distributions.

Two Methods Compared

A comparison of the two methods—the cross-indexing of response distributions and the backward-forward-association approach—is presented here. Words selected for which stimulus hierarchies were to be obtained included only reliable noun-primaries which are found in both the Russell-Jenkins and the Bilodeau-Howell norms. These included the following: chair, flower, grass, ground, insect, lamb, queen, sky, song, and water.

The cross-indexing procedure was performed using the Bilodeau-Howell norms, where a simple clerical search was made of response distributions to 180 stimulus words. For example, each time the word *sky* was found in a hierarchy, the stimulus was listed together with the probability of its evoking *sky*. At least minimal success was insured, since each word of interest was, from the beginning, the primary for some stimulus in those norms. The results of this search are presented in the left half of Table 1. It will be noted that a large number of strongly effective stimuli were located for the word *water*. Their probabilities are distributed over a wide range (.48 to .02) at fairly regular intervals. Several adequate stimuli for *flower* were also located. Most of the remaining hierarchies suffer from having too few or no strong stimuli. One should, however, remember that these hierarchies cannot be evaluated like response distributions. While all stimuli thus far investigated have produced response distributions of some sort, it is quite possible that some words have no effective stimulus hierarchy at all (e.g., *seahorse* may not be anyone's dominant response to any single word). Thus, a small stimulus hierarchy is not necessarily the result of inadequacies of the clerical approach.

Rogers (1965) used the backward-association approach to gather lists of potential stimuli for the same 10 response words used in the cross-indexing described above. Instructions to Ss for this procedure were mentioned earlier. Words that occurred most frequently (based on an N of 100) were then administered to a separate sample of 50 Ss by Snell (1966), using word-association procedures outlined by Bilodeau & Howell (1965).

The stimulus hierarchies obtained by means of this backward-forward-association method are presented in the right half of Table 1. These hierarchies, on the whole, contain more stimuli ($\overline{x} = 8.2$) than do the clerical distributions ($\overline{x} = 4.5$). Furthermore, these additional items are, on the average, equal in associative strength, the mean probability of response being .20 for each method. A word-by-word analysis reveals that for some words (e.g., *water*), either approach provides much useful information. For other words, such as *chair*, the backward-association approach yielded more and stronger stimuli. In no instance, among these 10 words, did the clerical technique seem more fruitful.

Did either approach produce words which were not already identifiable from looking at a word's response distribution? Of the 128 stimulus words in Table 1, 60 are found also in the response hierarchies of the words they elicit. Many of these 60 items are strong elicitors of the word of interest. Within the stimulus hierarchies obtained by means of the cross-indexing method, a +.39 correlation between frequency of forward

Table 1
Comparison between Stimulus Hierarchies Obtained for 10 Response Words using Two Different Methods of Assessment

Response word	Clerical cross-indexing method[a]		Backward-forward-association method[b]	
	Stimulus word	p[c]	Stimulus word	p[c]
Chair	table*	.50	rocking	.84
			table*	.56
			seat*	.46
			desk*	.30
			sit*	.16
			arm*	.12
			lawn	.06
			daisy	.02
Flower	bloom	.67	tulip	.70
	blossom	.65	bloom	.66
	stem	.39	daisy*	.62
	bud*	.38	petal*	.54
	plant	.37	garden*	.34
	bulb	.05	meadow	.04
	butterfly	.05	pot	.04
	yellow	.02	pretty*	.02
			wall	.02
Grass	green*	.31	green*	.40
	ground	.02	mower	.34
	leaf	.02	lawn*	.30
	plant	.02•	Bermuda	.28
			meadow	.20
			snake	.18
			cricket	.04
			seed*	.04
			cut*	.02
			plow	.02
Ground	earth*	.24	earth*	.16
	dirt*	.13	soil*	.16
			plow	.12
			dirt*	.08
			sky*	.02
Insect	butterfly	.22	cricket	.14
	moth	.16	bug*	.12
			DDT	.12
			roach*	.10
			ant*	.10
			mosquito*	.08
			bee*	.04
			gnat	.04
			fleece	.02
Lamb	mutton	.39	ewe	.20
	sheep*	.19	sheep*	.16
			fleece	.14
			wool*	.12
			baa	.08
Queen	king*	.66	king*	.52
	crown*	.09	royalty*	.20
			princess	.14
			crown*	.12
			daisy	.02
Sky	blue*	.22	stars*	.54
	ground	.12	cloud*	.46
	earth	.03	blue*	.34
	bright*	.02	earth	.12
	high*	.02	clear	.08
	king*	.02	high*	.08
			moon	.08
			heaven	.06
			air	.02
			sea*	.02
Song	music*	.16	sing*	.60
	whistle	.02	lyrics*	.50
			tune	.30
			music*	.12
			record*	.08
Water	thirsty	.48	thirsty	.64
	lake*	.48	lake*	.30
	drink*	.39	drink*	.26
	sea*	.31	sea*	.22
	river	.29	wet*	.14
	stream	.29	clear	.12
	ocean*	.25	glass	.12
	soap	.25	rain*	.12
	bath	.13	ship	.08
	food	.05	mud	.06
	blue	.04	air	.04
	ice*	.04	tower	.04
	hot	.03	Bermuda	.02
	cold*	.02	cold	.02
	milk	.02	garden	.02
	rough	.02	green	.02
			sweet	.02

[a] *Clerical cross-indexing of the Bilodeau-Howell (1965) norms (.01 responses excluded).*

[b] *Data adapted from Snell (1965).*

[c] *Probability with which stimulus evokes given response in word association.*

**Stimulus word found also in response hierarchy of word which it elicits.*

elicitation and frequency of backward association was found for such double-entry words. There are, of course, many members of a word's response distribution which will not in turn stimulate that word as a response. It appears, however, that to a fair degree, items high in a word's response hierarchy will tend also to be strong stimuli, if they elicit the word at all.

The average 50% overlap found between a word's stimuli and its responses is an inadequate indicator in that (a) the sample of stimuli obtained cannot be considered exhaustive for any given word, and (b) not all responses in the Bilodeau-Howell response hierarchies have been also used as stimulus words in a forward-association task. However imperfect the data in the present study may be, though, they do demonstrate (a) that sampling from stimulus hierarchies is feasible, and (b) that a word's stimulus hierarchy differs from its response hierarchy both in terms of items included and associative strengths.

One might wonder whether backward association alone would be sufficient to generate hierarchies. If a strong enough relationship existed between backward association and subsequent forward association, the second step could be dropped. Unfortunately, such does not obtain (r = +.36), possibly for several reasons. Rogers's Ss were asked to produce a stimulus for a word, not a stimulus which would selectively evoke that word as a primary, nor even a stimulus which would not evoke other competing responses. Hence, S may list *red* as a stimulus for *apple*; yet when asked to respond in discrete free association to *red*, he may reply with *blue*. Obviously, continued associations might be more sensitive indicators under these circumstances. *Apple* might then be elicited on the second, third, or later presentation of *red*. It might be desirable, then, to use highly trained rather than naive Ss to produce potential stimuli via backward association, thereby decreasing some of the noise in the system. A further problem exists, though, with respect to probability as a response measure. For both backward- and forward-association norms, probability of response is based on relative frequency of occurrence among events. An index of absolute strength of habit is badly needed, and would be much more useful.

Assessment of stimulus hierarchies, while feasible, exceeds by an order of magnitude the difficulty of obtaining norms on response distributions. None of the approaches presented in this paper is without strong disadvantages, and no technique presently available insures the location of all strong stimuli for a given word. The adequacy of information gained by any of the techniques listed will depend on the user's specific needs. It is, then, reasonable to wonder what need exists for data on stimulus hierarchies. One cannot foresee, when collecting norms of any sort, the extent of their eventful applications. Surely applications of the Russell-Jenkins (1954) norms and the Thorndike-Lorge (1944) word count have far exceeded author's expectations. At the moment, though, certain applications are already apparent in (a) pursuing Deese's (1961) attack on meaning, (b) paired-associates training through manipulation of associative S-R strength via E's choice of stimuli, response components held constant, (c) pursuing the knotty problem of verbal stimulus generalization, and (d) elucidation of rules for constituent synthesis in verbal behavior (a very nagging problem for most stimulus-response theorists) via use of multiple stimuli to elicit verbal responses, as set forth by Miller (1965) and Underwood (1966).

REFERENCES

BILODEAU, E. A., FOX, P. W., & BLICK, K. A. Stimulated verbal recall and analysis of sources of recall. *J. verbal Learn. verbal Behav.*, 1963, 2, 422-428.

BLICK, K. A. Cultural primaries as a source of interference in short-term verbal retention. *J. exp. Psychol.*, 1965, 69, 246-250.

COFER, C. N. On some factors in the organizational characteristics of free recall. *Amer. Psychologist*, 1965, 20, 261-272.

COLEMAN, E. B. The association hierarchy as an indicator of extra-experimental interference. *J. verbal Learn. verbal Behav.*, 1963, 2, 417-421.

DEESE, J. On the structure of associative meaning. *Psychol. Rev.*, 1962, 69, 161-175.

JENKINS, J. J. Mediated associations: Paradigms and situations. In C. N. Cofer and B. S. Musgrave (Eds.), *Verbal behavior and learning*. New York: McGraw-Hill, 1963. Pp. 210-245.

MILLER, G. A. Some preliminaries to psycholinguistics. *Amer. Psychologist*, 1965, 20, 15-20.

MINK, W. D. Semantic generalization as related to word association. Tech. Rep No. 17, 1957, Contract No. N8onr-66216, Office of Naval Research and University of Minnesota.

ROGERS, C. A., JR. Associative hierarchies, stimulus sampling, and differential recall. Unpublished doctoral dissertation, Tulane University, 1965.

ROSENBERG, S. Recall of sentences and function of syntactic and associative habit. *J. verbal Learn. verbal Behav.*, 1966, 5, 392-396.

RUSSELL, W. A., & JENKINS, J. J. The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff Word Association Test. Tech. Rep. No. 11, 1954, Contract No. N8onr-66216, Office of Naval Research and University of Minnesota.

SNELL, G. R. Associative stimulus hierarchies. Unpublished master's thesis, University of Arizona, 1966.

THORNDIKE, E. L., & LORGE, I. *The teacher's word book of 30,000 words*. New York: Teacher's College, 1944.

UNDERWOOD, B. J. Interference and forgetting. *Psychol. Rev.*, 1957, 64, 49-60.

UNDERWOOD, B. J. Motor-skills learning and verbal learning: Some observations. In E. A. Bilodeau (Ed.), *Acquisition of skill*. New York: Academic Press, 1966. Pp. 489-516.

UNDERWOOD, B. J., & POSTMAN, L. Extraexperimental sources of interference in forgetting. *Psychol. Rev.*, 1960, 67, 73-95.

WEBSTER'S seventh new collegiate dictionary. (7th ed.) Springfield, Mass.: G. & C. Merriam, 1967.

NOTE

1. Preparation of this manuscript was supported in part by NASA Grant NGR 03-002-091.

GSR amplitude instead of GSR magnitude: Caveat emptor![1]

H. D. KIMMEL, *DEPARTMENT OF PSYCHOLOGY, OHIO UNIVERSITY, Athens, Ohio 45701*

Size of GSR made by 300 Ss to 20 repetitions of a visual stimulus is presented as mean magnitude *and mean* amplitude, *illustrating that the* amplitude *method (averaging only those responses which are greater than zero) is susceptible to distortion resulting from a systematic elimination of Ss who initially make small responses. From trial to trial, the* amplitude *function comes more and more to be due to the behavior of Ss who initially make large responses. Even though the latter Ss' responses actually reduce across trials, the* amplitude *function rises.* Magnitude *(including zeroes) does not suffer from this distortion. It is suggested that the definition of a zero response is the source of the problem.*

In a recent discussion of methodological problems in classical conditioning, Gormezano (1966) has urged that researchers using the galvanic skin response (GSR) should use *amplitude* rather than *magnitude* of the response in describing their results. His argument is based on the proposition that "it is desirable to make frequency and extent measures as independent as possible" (p. 386). Magnitude measures are typically obtained by summing the reactions of all of the Ss on a particular occasion and dividing by the total number of Ss; amplitude measures differ in that the same sum is divided by only the number of Ss who actually made non-zero reactions on that occasion. On the basis of this distinction, Morrow & Keough (in press), Prokasy & Ebel (1964, 1967), and Prokasy, Fawcett, & Hall (1962) have adopted the procedure of reporting amplitude and magnitude measures separately.