

Psychological Bulletin

THE NATURE AND MEASUREMENT OF MEANING¹

CHARLES E. OSGOOD

*University of Illinois*²

The language process within an individual may be viewed as a more or less continuous interaction between two parallel systems of behavioral organization: sequences of central events ("ideas") and sequences of instrumental skills, vocalic, gestural, or orthographic, which constitute the communicative product. A communicator vocalizes, "It looks like rain today; I'd better not wash the car." This output is a sequence of skilled movements, complicated to be sure, but not different in kind from tying one's shoes. Even the smallest units of the product, phonetic elements like the initial "l"-sound of "looks," result from precisely patterned muscle movements. The organization of these movements into word-units represents skill sequences of relatively high predictability; certain longer period sequences involving syntactical order are also relatively predictable for a given language system. But execution of such sequences brings the communicator repeatedly to what may be called "choice-points"—points where the next skill sequence is not highly predictable from the objective communicative product itself. The dependence of "I'd better not wash the car" upon "looks like rain today," the *content* of the message, reflects determinants within the semantic system which effectively "load" the transitional probabilities at these choice-points.

It is the communicative product, the spoken or written words which follow one another in varying orders, that we typically observe. Since we are unable to specify the stimuli which evoke these communicating reactions—since it is "emitted" rather than "elicited" behavior in Skinner's terminology (97)—measurements in terms of rates of occurrence and transitional probabilities (dependence of one event in the stream upon others) are particularly appropriate (cf., Miller, 76). In-

¹ The research on which in part this report is based is being supported jointly by the University Research Board of the University of Illinois and the Social Science Research Council. Grateful acknowledgment is made to both institutions for their assistance.

² Institute of Communications Research and Department of Psychology.

terest may be restricted to the lawfulness of sequences in the observable communicative product itself, without regard to the semantic parallel. This is traditionally the field of the linguist, but even here it has proved necessary to make some assumptions about meaning (cf., Bloomfield, 4). On the other hand, one may be specifically interested in the semantic or ideational level. Since he is presently unable to observe this level of behavior directly, he must use observable characteristics of the communicative product as a basis for making inferences about what is going on at the semantic level. He may use sequential orderliness in the product to draw conclusions as to semantic orderliness in the speaker's or writer's mediation processes (i.e., which "ideas" tend to go together in his thinking with greater than chance probabilities). Or he may wish to study the ways in which central, semantic processes vary from concept to concept, from person to person, and so on. It is the problem of measuring meaning in this latter sense which will be discussed in the present paper.

Before inquiring into the measurement of the meaning of signs, for which there are no accepted, standardized techniques available, we may briefly mention certain fairly standard methods for measuring the comparative strength of verbal habits. Thorndike and his associates (102, 103) have made extensive *frequency-of-usage counts* of words in English; that this method gets at the comparative habit strengths of word skill sequences is shown by the fact that other measures of response strength, such as latency and probability within the individual (Thumb and Marbe, 106; Cason and Cason, 19), are correlated with frequency-of-usage. Zipf (117, 118, and elsewhere) has described innumerable instances of the lawfulness of such habit-strength measures. Whether samples be taken from Plautine Latin, newspaper English, or the English of James Joyce in his *Ulysses*, a fundamental regularity is found, such that frequency of occurrence of particular words bears a linear relation to their rank order in frequency, when plotted on double-log paper (Zipf's Law). Measurement of flexibility or diversity in communicative products is given by the *type-token ratio* (TTR): with each instance of any word counting as a token and each different word as a type, the greater the ratio of types to tokens the more varied is the content of a message. This measure can be applied comparatively to different forms of material, different kinds of individuals, and so forth (cf., Carroll, 16, 17; Johnson, 45; Chotlos, 20), provided the sizes of samples are constant. One may also count the ratios of adjectives to verbs (Boder, 5), the frequencies of different pronouns, intensives, and so forth (cf., Johnson, 45).

Although the above measures get at the comparative strengths of verbal skill sequences per se (i.e., without regard to meaning), this is not a necessary restriction. Frequency counts of this type can be applied to *semantic habit strengths* as well. Skinner (96) has shown that a similar lawfulness applies to the frequencies of "free" associations in the Kent-Rosanoff tests. When frequencies of particular associates to given stimulus words for a group of subjects are plotted against their rank order in frequency, a straight-line function on double-log paper results (Zipf's Law). In other words, associations at the semantic level appear to be organized in such a way that few have very high probability of occurrence and many have low probabilities of occurrence. Bousfield and his collaborators (7, 8, 9, 10, 11) have described a *sequential association method* for getting at comparative semantic habit strengths. When subjects associate successively from the same "pool," e. g., "names of four-legged animals," (a) the rate of successive associates shows a negatively accelerated curve, (b) varying in its constants with certain characteristics of materials and subjects, (c) the order of appearance of particular associates in individuals being predictable from the frequency of usage in the group, and (d) distortions in the function being related to particular transitional probabilities among associates, i.e., clustering. Useful though these measures are for many purposes, they do not get at meaning. The fact that "dog" has a higher probability of occurrence in sequential association than "otter" says nothing whatsoever about the differences in meaning of these two signs.

An extensive survey of the literature fails to uncover any generally accepted, standardized method for measuring meaning. Perhaps it is because of the philosophical haziness of this concept, perhaps because of the general belief that "meanings" are infinitely and uniquely variable, or perhaps because the word "meaning" as a construct in our language connotes mental stuff, more akin to "thought" and "soul" than to anything observable—for some combination of reasons there has been little attempt to devise methods here. Nevertheless, whether looked at from the viewpoints of philosophy or linguistics, from economic or sociological theory, or—interestingly enough—from within the core of psychological theories of individual behavior, the nature of meaning and change in meaning are found to be central issues. The proposals to be made in the latter portion of this paper are part of a program aimed at the development of objective methods of measuring meaning. Beyond obvious social implications, it is felt that this direction of research is a logical extension of scientific inquiry into an area generally considered immune to its attack.

THEORIES OF MEANING

Not all stimuli are signs. The shock which galvanizes a rat into vigorous escape movements usually does not stand for anything other than itself, nor does the pellet of food found at the end of a maze, nor a hammer in one's hand or a shoe on one's foot. The problem for any meaning theorist is to differentiate the conditions under which a pattern of stimulation is a sign of something else from those conditions where it is not. This certainly seems simple enough, yet it has troubled philosophers for centuries. By stating the problem somewhat formally, the chief differences between several conceptions of the sign-process can be made evident: let

\dot{S} = object = any pattern of stimulation which evokes reactions on the part of an organism, and

\boxed{S} = sign = any pattern of stimulation which is not this \dot{S} but yet evokes reactions relevant to \dot{S} —conditions under which this holds being the problem for theory.

The definition of \dot{S} is broad enough to include any pattern of stimulation which elicits any reaction from an organism. Although one usually thinks of "objects" as those things denoted by signs, actually any pattern of stimulation—a gust of northerly wind against the face, the sensations we call "belly-ache," the sensations of being rained upon—is an "object" at this level of discourse. One sign may be the "object" represented by another sign, as when the picture of an apple is called "DAX" in certain experiments. The definition of \boxed{S} is purposely left incomplete at this point, since it depends upon one's conception of the nature of the sign-process.

We may start a logical analysis of the problem with a self-evident fact: *the pattern of stimulation which is the sign is never identical with the pattern of stimulation which is the object.* The word "hammer" is not the same stimulus as is the object hammer. The former is a pattern of sound waves having characteristic oscillations in frequency and intensity; the latter, depending upon its mode of contact, may be a visual form having characteristic color and shape, a pattern of tactual and proprioceptive sensations, and so on. Similarly, the buzzer in a typical rat experiment is not identical as a form of stimulation with the shock which it comes to signify. Yet these signs—the word "hammer" and the buzzer—do elicit behaviors which are in some manner relevant to the objects they signify, a characteristic *not* shared with an infinite number of other stimulus patterns that are *not* signs of these objects. In simplest terms, therefore, the question is: *under what conditions does something*

which is not an object become a sign of that object? According to the way in which this question is answered we may distinguish several theories of meaning.

Mentalistic View

The classic interpretation derives directly from the natural philosophy of Western culture, in which the dualistic connotations of language dictate a correlation between two classes of events, material and nonmaterial. Since meanings are obviously "mental" events and the stimuli representing objects and signs are obviously "physical" events, any satisfying theory of meaning must specify interrelation between these levels of discourse. At the core of all mentalistic views, therefore, we find an "idea" as the essence of meaning; it is this mental event which links or relates the two different physical events, sign and object. The word "hammer" gives rise to the idea of that object in the mind; conversely, perception of the object hammer gives rise to the same idea, which can then be "expressed" in appropriate signs. In other words, *something which is not the object becomes a sign of that object when it gives rise to the idea associated with that object.* Probably the most sophisticated expression of this view is given by Ogden and Richards (82) in their book, *The Meaning of Meaning*. Most readers will recall their triangular diagram of the sign-process: the relation between symbol and referent (the base of their triangle) is not direct but inferred, mediated through mental "thought" or "interpretation" (the third corner of their triangle).

Substitution View

Naive application of Pavlovian conditioning principles by early behaviorists like Watson led to the theory that signs achieve their meanings simply by being conditioned to the same reactions originally made to objects. This, in essence, is the view one encounters in many introductory texts in general psychology. An object evokes certain behavior in an organism; if another pattern of stimulation is consistently paired with the original object, it becomes conditioned to the same responses and thus gets its meaning. The object is the unconditioned stimulus and the sign is the conditioned stimulus, the latter merely being substituted for the former. The definition of the sign-process here is that *whenever something which is not the object evokes in an organism the same reactions evoked by the object, it is a sign of that object.* The very simplicity of this theory highlights its inadequacy. Signs almost never evoke the *same* overt responses as do the objects they represent. The word FIRE has meaning to the reader without sending him into headlong flight. Nevertheless, this represents a first step toward a behavioral interpretation of the sign-process.

Meaning as "Set" or "Disposition"

In a monograph entitled *Foundation of the Theory of Signs* (77), Charles Morris, a semiotician working in the tradition established by Peirce and other American pragmatists, proposed a formula for the sign-process which avoids the pitfalls of substitution theory but seems to step backward toward the mentalistic view. In essence he states that signs achieve their meanings by eliciting reactions which "take account of" the objects signified. The sign "hammer" may evoke quite different responses from those evoked by the object signified, but these responses must have the character of being relevant to the object. The response made to the sign is called the "interpretant" which mediately takes account of the object signified. But it would seem that this process of "taking account of" is precisely what needs elucidation.

During the period intervening between this monograph and his recent book, *Signs, Language and Behavior* (78), Morris studied with two prominent behavior theorists, Tolman and Hull. The effects of this immersion in learning theory are evident in his book, which is a pioneer attempt to reduce semiotic to an objective behavioral basis. He states that "if anything, *A*, is a preparatory stimulus which in the absence of stimulus-objects initiating response-sequences of a certain behavior-family causes a disposition in some organism to respond under certain conditions by response-sequences of this behavior-family, then *A* is a sign" (p. 10). Reduced to its essentials and translated into our terms, this becomes: *any pattern of stimulation which is not the object becomes a sign of that object if it produces in an organism a "disposition" to make any of the responses previously elicited by that object.* There is no requirement that the overt reactions originally elicited by the object also be made to the sign; the sign merely creates a disposition or set to make such reactions, actual occurrence depending upon the concurrence of supporting conditions.

Beyond the danger that "dispositions" may serve as mere surrogates for "ideas" in this theory, there are certain other difficulties with the view as stated. For one thing, Morris seems to have revived the substitution notion. The sign is said to dispose the organism to make overt response-sequences of the *same* behavior-family originally elicited by the object. But is this necessarily the case? Is my response to the word "apple" (e.g., free-associating the word "peach") any part of the behavior-family elicited by the object apple? For another thing, Morris' formulation fails to differentiate sign-behavior from many instinctive reactions and from ordinary conditioning. To appreciate this difficulty will require a brief digression.

When a breach is made in a termite nest, the workers set up a distinctive pounding upon the floor of the tunnel and the warriors come charging to the spot, where they take up defensive positions. Is this pounding sound a sign to the warrior-termites that there is a breach in the nest? It happens that this

behavior is purely instinctive, and most students of sign-behavior believe that signs must achieve their signification through *learning*. But is learning a sufficient criterion? Are all stimuli that elicit learned reactions automatically signs? In developing any skill, such as tying the shoes, the proprioceptive stimuli produced by one response become conditioned to the succeeding response—but of what are these proprioceptive stimuli signs? With repeated experience on an electrified grill a rat will often learn to rear up on its hind legs and alternately lift them, this act apparently reducing the total intensity of pain—the painful stimulation is thus conditioned to a new response, but of what is the pain a sign?

If only some of the stimuli which elicit learned responses are signs, we must seek a reasonable distinction *within* the class of learned behaviors. We cannot draw a line between human and subhuman learning; the buzzer is operationally as much a sign of shock to the rat in avoidance-training experiments as are dark clouds a sign of rain to the professor—both stimuli elicit reactions appropriate, not to themselves, but to something other than themselves. Is voluntariness of response a criterion? Meaningful reactions may be just as involuntary as perceptions—try to observe a familiar word and avoid its meaning! Is it variability of response to the stimulus? Meaningful reactions may be just as stable and habitual as motor skills.

The Mediation Hypothesis

I shall try to show that the distinguishing condition of sign behavior is the presence or absence of a *representational mediation process* in association with the stimulus. This conception of sign behavior is based upon a general theory of learning rather than being concocted specifically to account for meaning as seen in human communication.³ The essence of the viewpoint can be given as follows:

1. *Stimulus-objects (S) elicit a complex pattern of reactions from the organism, these reactions varying in their dependence upon presence of the stimulus-object for their occurrence.* Electric shock galvanizes the rat into vigorous jumping, squeaking, and running activities, as well as autonomic "anxiety" reactions. Food objects elicit sequences of salivating, chewing, lip-smacking, and so forth. Components like salivating and "anxiety" are relatively independent of the food or shock stimulation respectively and hence can occur when such objects are not present.

2. *When stimuli other than the stimulus-object, but previously associated with it, are later presented without its support, they tend to elicit some reduced portion of the total behavior elicited by the stimulus-object.* This reduction process follows certain laws: (a) mediating reactions which interfere with goal-achievement tend to extinguish; (b) the more energy expenditure involved in making a particular reaction, the less likely it is to survive the reduction process; (c) there is evidence that certain reactions (e.g., autonomic) condition more readily than

³ This hypothesis, as an elaboration from Hullian theory (43), is described in my forthcoming book, *Method and Theory in Experimental Psychology*.

others (e.g., gross skeletal) and hence are more likely to become part of the mediation process—this may merely reflect factor (b) above.

3. *The fraction of the total object-elicited behavior which finally constitutes the stable mediation process elicited by a sign (\boxed{S}) will tend toward a minimum set by the discriminatory capacity of the organism.* This is because the sole function of such mediating reactions in behavior is to provide a distinctive pattern of self-stimulation (cf., Hull's conception of the "pure stimulus act").

4. *The self-stimulation produced by sign-elicited mediation processes becomes conditioned in varying strengths to the initial responses in hierarchies of instrumental skill sequences.* This mediated self-stimulation is assumed to provide the "way of perceiving" signs or their "meaning," as well as mediating instrumental skill sequences—behaviors to signs which take account of the objects represented.

Whereas Morris linked sign and object through partial identity of object-produced and disposition-produced behaviors, we have linked sign and object through partial identity of the "disposition" *itself* with the behavior elicited by the object. Words represent things because they produce some replica of the actual behavior toward these things, as a mediation process. This is the crucial identification, the mechanism that ties particular signs to particular stimulus-objects and not to others. Stating the proposition formally: *a pattern of stimulation which is not the object is a sign of the object if it evokes in an organism a mediating reaction, this (a) being some fractional part of the total behavior elicited by the object and (b) producing distinctive self-stimulation that mediates responses which would not occur without the previous association of nonobject and object patterns of stimulation.* This definition may be cumbersome, but all the limiting conditions seem necessary. The mediation process must include some part of the same behavior made to the object if the sign is to have its particularistic representing property. What we have done here, in a sense, is to make explicit what may be implicit in Morris' term "disposition." The second stipulation (b) adds the learning requirement—the response of warrior-termites to pounding on the tunnel floor is ruled out since it does not depend upon prior association of pounding with discovery of a breach in the nest.

Paradigm A in Figure 1 gives an abbreviated symbolic account of the development of a *sign*, according to the mediation hypothesis. Take for illustration the connotative meaning of the word SPIDER. The stimulus-object (\dot{S}), the visual pattern of hairy-legged insect body often encountered in a threat context provided by other humans, elicits a complex pattern of behavior (R_T), which in this case includes a heavy loading of autonomic "fear" activity. Portions of this total behavior to the spider-object become conditioned to the heard word, SPIDER. With repetitions of the sign sequence, the mediation process becomes reduced to some minimally effortful and minimally interfering replica—but still includes those autonomic reactions which confer a threatening

significance upon this sign. This mediating reaction (r_m) produces a distinctive pattern of self-stimulation (s_m) which may elicit a variety of overt behaviors (R_x)—shivering and saying “ugh,” running out of a room where a spider is said to be lurking, and even refusing a job in the South, which is said to abound in spiders.

The vast majority of signs used in ordinary communication are what we may term *assigns*—their meanings are literally “assigned” to

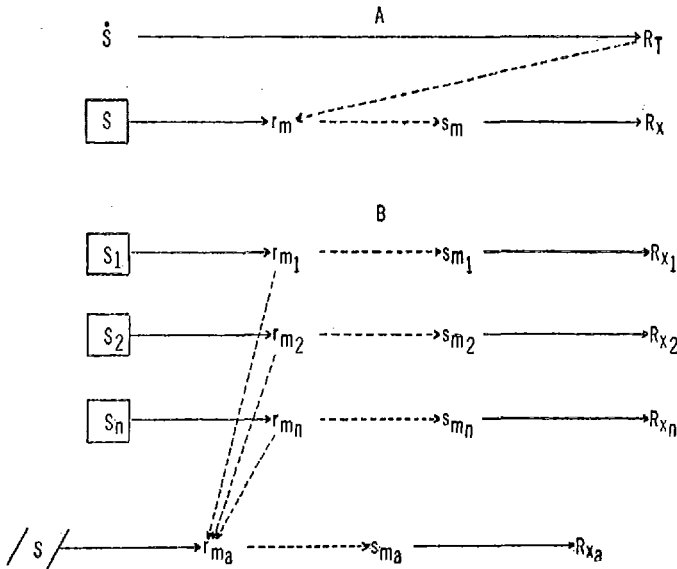


FIG. 1. SYMBOLIC ACCOUNT OF THE DEVELOPMENT OF SIGN PROCESSES:
 A. DEVELOPMENT OF A SIGN; B. DEVELOPMENT OF AN ASSIGN.

them via association with other signs rather than via direct association with the objects represented. The word ZEBRA is understood by most six-year-olds, yet few of them have ever encountered zebra-objects themselves. They have seen pictures of them, been told they have stripes, run like horses, and are usually found wild. As indicated in Figure 1 (B), this new stimulus pattern, ZEBRA, “picks up” by the mechanisms already described portions of the mediating reactions already elicited by the primary signs. In learning to read, for example, the “little black bugs” on the printed page are definitely assigns; these visual patterns are seldom directly associated with the objects signified, but rather with auditory signs (created by the child and teacher as they verbalize). Obviously, the more quickly the child can learn to make the right noises to these visual stimuli (the modern phonetic approach to reading), the more quickly these new, visual assigns will acquire signifi-

cance. The child already has meanings for HOUSE, DOG, and even TYPEWRITER as *heard* stimulus patterns, but these mediation processes must be assigned to *seen* stimulus patterns.

It is apparent from the foregoing that the meanings which different individuals have for the same signs will vary with their behaviors toward the objects represented. This is because the composition of the mediation process, which *is* the meaning of a sign, is entirely dependent upon the composition of the total behavior occurring while the sign-process is being established. This indicates that to change the meaning of signs we must change behavior with respect to objects (keeping in mind that the "objects" for assigns are other signs). On the other hand, meanings are quite independent of the stimulus characteristics of the signs themselves, a point repeatedly stressed by linguists. According to the present theory, there is nothing sacred about the particular mouth-noises we use in communication any more than there is about the buzzer that becomes a sign of shock to the rat—a flash of light or a blast of air would serve as well. Of course, in human communication (in contrast to sign-behavior in the rat) it is necessary that the users of signs be able to produce as well as receive them.

EVALUATION OF EXISTING TECHNIQUES OF MEASUREMENT

Physiological Methods

Meaning has been identified with representational mediation processes. Although no conclusive evidence as to the nature of these processes can be given at this time—whether they require participation of peripheral motor and glandular mechanisms or are sufficiently characterized as central phenomena—it is convenient to conceive of them as implicit response processes which produce self-stimulation. The meager evidence available certainly does not refute this view. In any case, the investigator is encouraged to see whether or not any physiological measures display covariance with changes in meaning. Correlates of this sort would be direct indices of meaning.

Action potentials in striate musculature. Working under the impetus given by the Würzburg School of imageless thought, introspective psychologists of another generation tried to tease out the sensory content of ideas. Students like Marbe (69), Clark (21), Washburn (110), Comstock (25) and Crosland (26) agreed in finding kinaesthetic sensations present as a residue when everything but "meaningful thought" was excluded. But here the limitations inherent in the introspective method made themselves apparent: pressed to the limits of critical self-analysis, the trained human observer reported vague muscular and organic sensations as being present during thought—but did these sensations constitute thoughts and meanings themselves or were they merely a back-

ground of bodily tonus? The method did not permit this kind of discrimination.

Experimentalists picked up the problem at precisely the point where the introspectionists had per force dropped it—was it possible that sufficiently sensitive apparatus could record the minimal motor events that eluded conscious analysis? Watson's (111) statement about thought being implicit speech was the "open sesame" for a small host of gadgeteers (Wyczoikowski, 116; Reed, 89; Clark, 22; Thorson, 104) who filled subjects' mouths with an astounding variety of mechanical devices and then had them both think and mumble unusual items like "psychology." That they found little or no correspondence between the movements made during thought and speech is not particularly surprising. The "thought" movements were results of irritation in all probability, and (contrary to Watson) there is no a priori reason to expect relevant mediation processes to be restricted to the organs of speech.

Apparatus of sufficient sensitivity came with the development of electrical methods of recording and amplifying action. Electrodes placed near the motor end-plates of efferent fibers pick up minute changes in potential which cause deflections in a sensitive string galvanometer. These signals are fed through a vacuum tube amplifier, magnified thousands of times, and recorded on photographic film. Using this method, Jacobsen (44) and Max (72, 73) obtained suggestive correlations between introspectively specifiable events and objectively recorded muscular events. Jacobsen showed, for example, that when a subject, trained in techniques of progressive relaxation, *imagines* lifting his right arm, distinctive potential changes are recorded from muscles in that member but not elsewhere. Max, using deaf-mutes as an ingenious control, was able to show that these individuals display more frequent and larger potential changes in the muscles controlling their fingers while solving mental problems and while dreaming than normal individuals; he also described a negative correlation between the magnitude of such implicit activity and intelligence scores, i.e., more intelligent subjects showed less overt activity.

Are these recordable motor events the mediating reactions (r_m) in the mediation processes we have specified as essential in sign behavior? Does this technique provide an index for the presence or absence of meaning, its degree, and quality? This is an attractive possibility, but the index is a crude one. There is no way of "reading" the meaning of a sign to a subject from the recorded activity. The ultimate criterion of meaning is still introspection of the subject—he verbalizes meaning while the experimenter scurries about his periphery trying to pick it up on instruments, and when activity fails to appear at predicted locus *a* it is assumed it must be occurring at some other locus *b*. No satisfactory demonstration of the *necessity* of the motor component has been offered; meaning might be present without measurable motor activity which,

when it does occur, is simply a specific overflow of excitation into motor pathways—an epiphenomenon. The same criticism applies with equal force to the other physiological correlates that have been studied.

Salivary reaction. Another pioneer investigation into the organic correlates of meaning was that by Razran (87), serving as his own subject. Meaningfulness of a series of signs was the independent variable, the stimuli being words for "saliva" in languages with which Razran had varying familiarity. Amount of salivary secretion was the dependent variable—following presentation of each stimulus, a dental cotton-roll was placed in his mouth for two minutes and its weight determined immediately afterward. As "meaningless" controls he used the Gaelic word for saliva, the nonsense syllables QER SUH, and periods of "blank consciousness." Salivation was greatest in his childhood tongue (Russian), next in his most proficient one (English), and less in three slightly known languages (French, Spanish, and Polish). The control conditions showed no differences among themselves, despite the fact that Razran "knew" the Gaelic word stood for saliva. This experiment demonstrates a relation between amount of salivation and degree of meaningfulness of signs to a sophisticated subject. We have here another feasible index of some aspect of meaning, albeit a very limited one.

The galvanic skin response. The GSR is one of several indices of autonomic activity, and to the extent that meanings include emotional components this measure should be useful. There are a large number of studies using the GSR that are remotely related to this problem: GSR is readily elicited by any warning or preparatory stimulus that precedes shock (Darrow and Heath, 28; Switzer, 101; Mowrer, 80), and it, therefore, may serve as an indication that the preparatory stimulus has become a sign of shock. It has been used to index the intensity of pleasant or unpleasant connotations of words and experiences (Jones, 46; Lynch, 63). In connection with free association, GSR has been found to be a good indicator of the emotional effect of stimulus words (Jones and Wechsler, 47). It is unfortunate, therefore, that the two most pertinent experiments in this area leave much to be desired in the way of methodological finesse.

Mason (70) asked this question: do changes in GSR accompany changes in meaning? What were called three types of change in meaning were studied: "certainty of meaning" (the expression of certainty by the subject as to the correctness of his recall in learning a list of nonsense syllables), "discovery of meaning" (the point in a series of readings of a trick sentence without punctuation at which the subject achieved insight), and "loss of meaning" (where the subject pressed a foot-pedal whenever the continuously vocalized word "tangerine" seemed to lose its meaning). Although the procedures and results of this experiment are reported in great detail, no tests of significance were employed and none of the necessary controls was introduced. In the

first two experiments, for example, it is impossible to determine whether the deflections in GSR should be attributed to change in meaning or simply to vacillations in emotional stress (the latter seems quite likely). In the experiment on loss of meaning, no check was made to see what effect simply pressing the indicator pedal and seeing a signal light come on might have had upon GSR.

Bingham (3) measured psychogalvanic reaction to 72 words "selected from the educational and philosophical writings of John Dewey and Rabindranath Tagore as the most frequently occurring words in samples of 8,000 from each writer." After the galvanic measurements were made, the 50 undergraduate subjects rated the words on a three-point scale in terms of their personal meaningfulness, significance, and importance (combined into a single scale, MSI). These untrained subjects further introspected on the sensory content in the meaning of each of these words, being requested to report any visual, auditory, kinesthetic, etc., imagery or sensations. Words having the highest MSI ratings yielded the greatest average change in skin resistance. High MSI words also had more "organic" sensory content, according to the introspections (the high "organic" content words were *intellectual, freedom, God, truth, and love*, in this order!). Much remains to be done in these directions.

Learning Methods

There are many learning studies employing meaningful materials, but rarely is meaning itself the experimental variable. Even where meaning has been deliberately varied, interest has generally centered on the effect upon learning rather than upon the use of learning as an index of meaning. Only the more relevant studies will be considered here.

Semantic generalization. When a reaction conditioned to one stimulus transfers to another, and the amount of transfer varies directly with the similarity between the two stimuli, we speak of stimulus generalization. The operations whereby semantic generalization is demonstrated are the same—except that the necessary similarities lie in meaning rather than in objective physical characteristics of the eliciting stimuli. There is no physical similarity between the word "blue" and light of 450 m μ yet generalization between such stimulations is easily demonstrated—we infer that the common overt reaction is mediated by some common implicit process. In the experiments to be summarized here, the precise nature of the overt reaction is unimportant—all the standard CR's have been used, salivation, GSR, finger retraction, pupillary reflex, and so on. Much of the research in this area is contributed by Russian investigators and, unfortunately, the available reports are mostly in the form of brief abstracts.

1. *From object to sign.* Kapustnik (49) set up conditioned salivary reactions to visual and auditory stimuli, transfer to verbal signs for the original cues being tested. Kotliarevsky (59) employed a cardio-vasomotor reflex: following conditioning to the sound of a bell, response to the word "bell" was tested. Metzner (75) reports a similar experiment with the pupillary reflex. In all these cases significant amounts of generalization were obtained. Traugott (107) was able to demonstrate the generalization of conditioned inhibition from blue light to the words, "blue, blue"—in fact, these words (quite different as physical stimuli) showed greater generalization than actual red light. This investigator was also able to show that the inhibitory effect of the words "blue, blue" transferred to *other* conditioned reactions more broadly than did the effect of blue light itself, a finding which fits well with our notions as to the abstractability of sign-processes. Traugott and Fadeyeva (108) combined conditioning and free association techniques: with excitatory CR's set up to a bell-whistle-light pattern and inhibitory CR's to a whistle-touch pattern, free associations to the verbal signs of these stimuli were recorded along with the latencies of these associations. Association to words representing conditioned excitors were made more rapidly than those to words representing conditioned inhibitors and, interestingly enough, after extinction of the excitatory CR's associations became slower and generically older, i.e., associations which had referred to the experimental situation now referred to pre-experimental situations.

2. *From sign to sign.* When a response is conditioned to one sign (e.g., the word TREE) and generalization to other signs (e.g., BUSH or the picture of a tree or bush) is measured, the essential role of meaningful mediation is merely more obvious than in the preceding situation. Razran (88) flashed single words and short sentences on a screen while six adults were eating—conditioned salivation developed rapidly. In a second session, different words and sentences were used to measure generalization. For the single words generalization was found to be greater for semantically related words (e.g., STYLE to FASHION) than for phonetically related words (e.g., STYLE to STILE), a result in keeping with Traugott's results discussed above and data more recently obtained by Riess (90) using the GSR as a measure. A slightly discordant note is contributed by Keller (55). After conditioning the GSR to a *picture* of a boy-scout hat, tests for generalization were made for a picture of a fireman's hat and the the printed word HAT, neutral control items being pictures of a duck and baseball and the words DUCK and BALL. While significant generalization to the picture of a fireman's hat occurred, no transfer to the word HAT was obtained. Keller argues reasonably that if the generalization between the two pictured hats was based on a common mediating response, thinking or subvocalizing "hat," the printed word should have shown the same effect. In one of the most interesting studies of this type, Riess (91) has related semantic generalization to stages in genetic development. Four groups of subjects, varying in mean age from 7:9 years to 18:6 years of age, were trained to give the GSR to selected verbal stimuli. Tests for transfer to synonyms, antonyms, and homonyms of the original words were run. The generalization results indicate that meaningful or semantic similarities (synonym and antonym relations) increase in importance as the individual matures while the importance of physical similarities (homonym relation) decreases.

3. There are other semantic relations that could be studied with similar

techniques, but there is little evidence available. Generalization between *hierarchical levels of signs*, e.g., between DOG and ANIMAL, has been studied by Goodwin, Long, and Welch (36), but there are certain difficulties with the research design (see below). Generalization would also be expected to occur *from sign to object*, even though this seems to reverse the sequence followed in the development of meaning. Kapustnik (49) found that salivary reactions conditioned to verbal signs transfer to the stimuli signified. This is the only directly relevant study I have uncovered, but observations on much of social behavior fit the paradigm. The prejudicial reactions associated with "Wop" and "Jap" on a verbal stereotype level certainly tend to transfer to the social objects represented, once they are encountered. Finally, mention should be made of generalization *from object to object via semantic mediation*. The reverential care with which the adolescent handles a certain handkerchief, a certain lock of hair, a certain lipstick-printed napkin has nothing to do with physical similarities among these objects themselves. Similarly, an "inferiority complex" may render a wide range of physically dissimilar social objects and situations equivalent in meaning and hence reaction to a particular individual (cf., G. W. Allport's [1] trait hypothesis).

Cofer and Foley have related the various studies of semantic generalization to the theoretical mediation process. They state that semantic generalization "thus presupposes and depends upon the preexperimental formation of conditioned responses or associations, i.e., *the gradient of generalization is a gradient along a dimension of conditioned stimulus functions*. The stimuli need be similar only in so far as they have previously been conditioned to the same (or similar) response" (23, p. 520). The pre-experimental formation of conditioned responses to which they refer is a special case of the formation of mediation processes, as discussed earlier in this paper. It is probably necessary to assume that primary generalization occurs both among mediating reactions and the stimuli they produce, in order to account for the fact that *gradients* of semantic generalization are correlated with *degrees* of meaningful similarity.

Transfer and interference in learning. The experiments contributed by Cofer and Foley in support of their hypotheses fit the standard transfer design. The general procedure was as follows: first a single repetition on a *buffer list* of numbers (spelled out) was given as warm-up; then an *equating list* of proper names was presented once and scored for recall immediately, subjects being assigned to various experimental conditions on this basis; each subject was then given four unscored repetitions on either a *reinforcement list* or a *control list*; finally, all subjects were tested for recall of a *test list* of words immediately after a single presentation. Experimental, control and test lists used in the first experiment (Foley and Cofer, 32) are given in Table 1. All words on a given reinforcement list bear the same relation to the test list, either some degree of synonymy or some degree of homonymy.

Although all experimental conditions yielded better recall scores than the control condition, there are several curious points about these results. In the first place, the best homonym list shows more "generalization" to the test list than the best synonym list—this is in flat contra-

diction to most other findings (cf., Traugott, Razran, Riess above). Secondly, the difference in recall between the two homonym lists is clearly greater than that between the two synonym lists, yet the former are obviously equivalent (both being identical in sound to the test words) while the latter are definitely not equivalent (the words on I are close synonyms of the test words but those on II bear no relation whatever to the test words, viz., sent-killed, vein-help, pear-result, sow-factory). Fortunately, the difference between "Synonym II" and the control condition was not significant.

TABLE 1
MATERIALS FOR AN EXPERIMENT ON MEDIATED TRANSFER
(FOLEY AND COFER, 32)

Control List	Reinforcement Lists				Test List
	Homonym I	Homonym II	Synonym I	Synonym II	
palm	cent	scent	dispatched	killed	sent
set	vain	vane	vessel	ship	vein
reed	pare	pair	fruit	result	pear
very	sew	so	plant	factory	sōw
numb	rain	rein	rule	principle	reign
me	seas	seize	looks	appearance	sees
day	write	rite	just	barely	right
snap	noes	nose	apprehends	arrests	knows
rope	meat	mete	join	enlist	meet
spire	dō	doe	batter	bruise	dough
<i>Mean Number of Words on Test List Recalled Following Above</i>					
4.80	6.72	5.64	5.88	5.24	

In searching for an explanation of these points, a flaw in design was discovered which renders this entire technique suspect. Since all the words on a given reinforcement list bore the same relation to those on the test list, all a subject had to do was to "catch on" to this abstract relation and then proceed to manufacture the test list rather than recall it. Given four trials on a list of only 10 meaningful words, the subject presumably masters most of it. If, on the single presentation of the test list, he now notes that the new words are homonyms of the old, by merely recalling *cent*, *vain*, and *pare* he can do a pretty good job of manufacturing (and checking by recognition) *sent*, *vein*, and *pear*. He can do the same thing with synonym lists, but here he will make more errors since there are more alternatives. The same loophole in design is even more apparent in later studies in this series. In a study on *antonym gradients*, for example, Cofer, Janis, and Rowell (24) themselves point out that 19 of 28 subjects reported that they recognized the opposition relation. Foley and Mathews (34) and Goodwin, Long and

Welch (36) report experiments using the same method, and their results are of dubious value for the same reasons.

Two experiments on interference in verbal learning by Osgood (83, 84) were explicitly designed to test certain hypotheses regarding the nature of meaning. The following hypotheses were set up: (1) Words of opposed meaning are so because they elicit reciprocally antagonistic mediating processes. (2) Repeated reinforcements of the association between a new stimulus and a particular mediating reaction produce a negatively accelerated increase in the excitatory tendency associating this stimulus with this reaction *and simultaneously an equal inhibitory tendency associating this stimulus with the reciprocally antagonistic reaction*. In other words, in learning to make a reaction to a stimulus the organism is simultaneously learning *not* to make the directly antagonistic reaction to that stimulus. (3) Both excitatory and inhibitory tendencies generalize in the usual fashion among both mediating reactions and the stimuli they produce.

With nonsense letter-pairs as constant stimuli and meaningful adjectives as varied responses in the standard retroactive interference paradigm ($A-B$; $A-K$; $A-B$), it was predicted that for both transfer and retroaction tests interpolated responses *similar* in meaning to the original responses should show the least interference (generalization of excitatory tendency) and responses *opposed* in meaning should show the most interference (generalization of inhibitory tendency), as compared with an intermediary neutral condition. The total design was such that each subject learned an equal number of items in each meaningful relation, thus avoiding the type of set that troubled the Foley and Cofer studies. The results for both transfer and retroaction situations were essentially those predicted (83). The second study (84) offered further evidence that a special form of reciprocal inhibition is operating in the successive learning of opposed meanings for the same sign. Different groups of subjects were given varying degrees of learning on the interpolated materials (cf. the design used by Melton and Irwin, 74) and only similar and opposed meaningful responses were compared. The learning of opposed responses was characterized by longer latencies of reaction and more frequent blanks (failures of response), both increasing with the degree of interpolated learning. Since both these phenomena are characteristic of weakened habit strength, they follow from the hypotheses.

These findings point to the following general conclusion: *when a sign or assign is conditioned to a mediator, it will also tend to elicit other mediators in proportion to their similarity to the original reaction; it will tend to inhibit other mediators in proportion to the directness of their antagonism to the original reaction*. In everyday language, this indicates that signs which develop a certain meaning through direct training will readily elicit similar meanings but resist being associated with opposed meanings. If the sign RUSSIAN means *bad* to the conservative college student he easily accepts substitution of *dirty*, *unfair*, and *cruel*, but it is

difficult for him to think of Russians as *clean*, *fair*, and *kind* (cf., Stagner and Osgood, 100).

Perception Methods

There is an intimate relation between perceptual and meaningful phenomena (14, 105). It is borne out by the confusions psychologists display in using these terms. In one of Maier's (67) ingenious insight situations, for example, the crux of the problem lay in whether or not the human subject could shift from utilizing the handle of an ordinary lab clamp as something to tighten (original use) to something to hang one's hat on (use which would solve the problem). We could say that this handle must be "perceived differently" or the "field restructured perceptually" (cf., Köhler, 58; Wertheimer, 112), or that it must be "given a new functional value" (cf., Duncker, 30), or that it must "acquire a new meaning or significance as a stimulus." The voluminous literature on memory for forms has been interpreted both as demonstrating perceptual dynamics (cf., Koffka, 57) and semantic dynamics (cf., Bartlett, 2)—witness particularly the experiment by Carmichael, Hogan, and Walter (15) in which the deliberate introduction of different meaningful words in association with the same abstract forms markedly influenced the way they were recalled.

As was the case with the learning approach, there are few experiments in which meaning has been deliberately introduced as a variable. Remotely relevant are a group of studies in which the effect of hunger upon perception in ambiguous situations has been measured (Sanford, 93, 94; Levine, Chein, and Murphy, 62; McClelland and Atkinson, 64). The way of perceiving the ambiguous stimuli was clearly modulated by the presence or absence of this motive state. Postman and Bruner (86) have studied the effect of a different motive state, frustration, upon the perception of tachistoscopically presented sentences. Most significant from our present point of reference was the marked increase in "aggressive" and "escape" words as misperceptions following frustration. Generalizing, we might say that the internal *motive state* of the individual, as part of the total stimulus context, changes the probabilities of occurrence (availability) of alternate mediating processes for the same external stimulus. Whether the known *value* of an object, as one dimension of its meaning, can influence the way it is perceived is a moot question at this time. Bruner and Goodman (12), with apparent size of coins as the perceptual characteristic measured and rich or poor 10-year-olds as subjects, obtained what may be interpreted as positive results; Carter and Schooler (18), under generally similar conditions, failed to substantiate the earlier conclusions. Mausner and Siegel (71),

using recognition-time as a measure and stamps of varying value as stimuli, also report negative conclusions. Most relevant to the problem of measuring meaning are the following experiments.

Bruner and Postman (13) compared the apparent sizes of a dollar sign (positive symbol), swastika (negative symbol), and an abstract geometrical design (neutral control), as estimated by manipulating a spot of light when plastic discs of identical size bearing these symbols were held in the subject's hands. Both dollar sign and swastika showed significant *overestimation*. According to the investigators, two dynamic processes were operative: (a) perceptual enhancement due to the positive value of the dollar sign and (b) perceptual accentuation of apparent size due to the swastika alerting the organism to danger or threat—a single process of enhancement in size due to distinctiveness might be more parsimonious, and there are many other possible hypotheses.

Postman, Bruner, and McGinnies (86) hypothesized that personal values (as defined by scores on the Allport-Vernon test) are among the behavioral or attitudinal determinants of perception. Twenty-five subjects were shown 36 words, one at a time, in a modified Dodge tachistoscope, these words being presented in random orders and being chosen to represent the six Allport-Vernon values. The usual method for obtaining recognition thresholds was used—gradually increasing the flash exposure time until the subject correctly identifies the word. Take for illustration a subject with high social values and low theoretical: according to the investigators, his threshold for perceiving words like "loving" and "devoted" should be lowered by *selective sensitization*; in the presolution period he should misperceive words covaluant with the correct word because of *value resonance*; and his threshold for words like "verify" and "research" should be raised because of *perceptual defense*. The results were consistent with the general thesis, but the single principle of selective sensitization seems sufficient to explain them.

McGinnies (65) inquired more penetratingly into the matter of "perceptual defense," conjecturing that autonomic reactions are aroused prior to conscious awareness of the meaning of a threatening word and hinder its perception. A list of 11 neutral words and seven emotionally charged words were presented tachistoscopically and recognition thresholds determined in the usual manner; GSR was recorded as a measure of emotional disturbance. Taboo words were found to require longer exposures for recognition and their prerecognition presentations were accompanied by significantly stronger emotional reactions. When asked if they had reported their perceptions promptly and accurately, all undergraduate subjects said they had.

Both this study and the preceding one on values have come in for their share of criticism. Howes and Solomon (40) argue that it is unnecessary to appeal to "selective sensitization" in the former case and "perceptual defense" in the latter, since recognition thresholds for

words have been shown to vary with their frequency of usage or familiarity (Howes and Solomon, 41) and this would provide a parsimonious explanation of both sets of findings. In the case of Allport-Vernon value systems, it seems reasonable to suppose that people with high theoretical values (and hence presumably with more courses, books, etc., in scientific fields) will have had more frequent visual contact with words like "logical" and "research"; in the case of taboo words, it also seems reasonable to suppose that the frequency of visual contact with words like "whore," "penis," and "bitch" is much less than with such control words as "child," "clear," and "dance." McGinnies' argument that these taboo words are much more frequent in ordinary conversations seems to be largely beside the point, since the test conditions were visual. The other main point raised by Howes and Solomon is that the emotional reactions *accompanied* recognition of the taboo words, appearing to precede simply because subjects inhibited reporting them—particularly since a member of the opposite sex was always present. These critics draw a delightful picture of what might have been going on in the subjects' minds during the McGinnies experiment. In an answer to these criticisms, McGinnies (66) draws a different picture. This is another issue yet to be resolved by further research. In this connection, a recent study by Lazarus and McCleary (61) reports that subjects may show heightened GSR to nonsense syllables previously associated with shock even when exposures are stopped prior to actual recognition. But here again one questions the validity of the demonstration—how can one react emotionally to the meaning of a sign before its significance has been appreciated? It may be that we will be forced to accept some conception of "unconscious" and "conscious" levels of perception or meaning.

Skinner (95) devised a "verbal summator" technique for studying language behavior which resembles these perception methods. Samples of meaningless speech sounds, obtained by permuting and combining elemental phonemes—a sort of verbal inkblot—are repeated until the subject himself perceives some meaningful form. According to Skinner, the verbal summator "evokes latent verbal responses through summation with imitative responses to skeletal samples of speech." That this method gets at the comparative strengths of verbal habits is indicated by the fact that the same double-log function of frequency to rank (Zipf's Law) appears when a large sample of such responses are analyzed. Estes (31) has described a visual form of summator which presents skeletonized verbal materials tachistoscopically.

Association Methods

Freud would have been the first to point out that the associations produced when a patient "allows one idea to lead to another" are in no sense "free" or random, but rather are semantically determined. An-

other analyst, Jung (48) used a more formal association method to get at the meanings of words to individuals. Lists of verbal stimuli calculated to touch off complexes were imbedded among neutral words. Among the indices of "unusual" responses was the rareness of the association itself. In order to judge the commonness or rareness of particular associations, it was necessary to know the comparative frequencies with which various responses to a given stimulus word occur in a representative sample of the population. Kent and Rosanoff (56) obtained responses to 100 common English nouns and adjectives from 1,000 subjects; their sample must have been fairly representative because the occasional re-checks that have been made show rather surprising agreement. Given norms like these, the unusualness of a subject's responses can be indexed by the frequency of occurrence of that response in the populations; *sharp* can be expected as a response to NEEDLE 152 times per 1,000 (15.2%), but *weapon* occurs only once (00.1%).

The gross majority of word associations are semantically determined, i.e., result from the mediation process set in motion by the verbal stimulus as a sign. All such associates are similar in some way to the stimulus word, either similar in meaning (NEEDLE-*pin*), which would include hierarchial relations (NEEDLE-*tool*), or in terms of commonness of context (NEEDLE-*thread*). The venerable associationistic principles of similarity and contiguity will be recognized here. After an intensive analysis of ways of classifying associates, Karwoski and Berthold (50) conclude that nearly all responses can be categorized as either some form of similarity or contrast. What about contrast responses? The single most frequent associate is often the direct opposite (LIGHT-*dark*; MAN-*woman*). For a number of reasons, this writer believes such contrast responses are *not* semantically determined at all, but rather reflect overlearning of verbal skill sequences, quite akin to FOOT-*ball*, APPLE-*cart*, and WASTE-*basket*. The tendency to free associate opposites increases with age, children readily giving similar and contextual responses but rarely opposites (cf., Woodworth, 114, p. 346). Furthermore, rather than being distributed among many varied but roughly equivalent words, as is the case with similar associates, the opposition tendency is largely restricted to a single word, the direct opposite (cf., Kent and Rosanoff tables). Karwoski and Schachter (54) report this same effect with opposites and add the fact that opposites are given with significantly shorter reaction-times than similars.

One of the more interesting applications of the association method has been in differentiating responses to sign, symbol, and object levels of stimulation. In an early study on this, Dorcus (29) compared associations to color words (signs) and actual bits of colored paper (objects).

Whereas co-ordinate and contrast responses were most common to color signs (WHITE-black, RED-blue), the names of contextually related objects were most commonly given to color objects (BLUE PAPER-ribbons; RED PAPER-fingernails). More recently Karwoski, Gramlich, and Arnott (51) have obtained associations to visually perceived actual objects, pictures of these objects, and verbal labels for these objects. The stimulus materials were such everyday things as *pipe*, *leaf*, *dollar*, and *pistol*. Where differences appeared, the dividing line was typically between the verbal level and the other two modes. For example, the most common response to the word FORK was, of course, *knife*—on both picture and objects levels the most common response was *eat*. Reaction times for the verbal level were also shorter.

Related to associational procedures are the effects of *context* upon meaning. It is a matter of common observation that a man's moods, emotions, and motives influence the character of his verbalizations. Bousfield and Barry (9) and Bousfield (8) found that subjects' rated moods (on a scale from "feeling well as possible" to "feeling as badly as possible") correlated with their rates of production of pleasant vs. unpleasant associates. The relatively stable attitudes of an individual also exert a contextual effect upon associations. Foley and MacMillan (33) have shown that associates to 40 ambiguous words (like *binding*, *administer*, *discharge*) are clearly influenced by the occupational status of subjects, as law students, medical students, or nonprofessional students. Perhaps because of the obviousness of the matter, no research seems to have been done upon the effect of the external, situational context upon meaning. This context includes the facial expressions and gestures of speakers, the objects present, and the activities underway, and so on (cf., Malinowski's [68] enlightening discussion of this in relation to decoding the language of another culture). Many slips of the tongue completely escape notice, simply because the situational context "carries" the intended meaning.

Howes and Osgood⁴ have given attention to the manner in which the meaning of a particular sign is affected by the pattern of verbal materials within which it is imbedded. A sequence of four spoken words made up each item, the first three serving as the context and the fourth, spoken with greater emphasis, serving as the actual stimulus for word-association by the subjects. One experiment was designed to get at the effect of varying the *density* of contextual items having a common semantic direction: Group A heard three contextual stimuli of very similar meaning (e.g., *sinister*, *devil*, *evil*-DARK), Group B had one

⁴ HOWES, D. H., & OSGOOD, C. E. Studies on the combination of associative probabilities in linguistic contexts (in preparation).

neutral word added (*eat, devil, evil-DARK*), Group C two neutral words (*eat, basic, evil-DARK*), and Group D, as control, had all neutral words in the context. When the frequencies of response-words related to the particular context (e.g., *thief, mystery, dead*, etc.) were plotted as a function of the density of influence in the context, number of influenced associates turned out to be a simple multiple of the number of relevant words in the context. In a second experiment, three influencing words of relatively independent meaning were used as "context," and the question was how *temporal proximity* of contextual stimuli affects the meaning of the eliciting sign. With an item like *feminine, strong, young-MAN*, the responses clearly relevant to each contextual stimulus could be isolated (*woman, girl* vs. *hard, work* vs. *boy, child*, for examples). Then the frequency of occurrence of such related responses was plotted as a function of the order of presentation of their contextual stimuli (e.g., frequencies of *woman* as a response when "feminine" is in third position, nearest MAN, second position, and first position, most remote from MAN). The results indicate that degree of influence of a contextual stimulus upon the meaning of a sign is a sharply negatively accelerated function of the temporal interval between them. In other words, the influence of one word upon another falls off rapidly as the amount of intervening material increases.

Scaling Methods

Considering the number of traits, abilities, and attitudes that psychologists have attempted to measure by scaling methods, it is significant that there has been practically no attempt to measure meaning this way. Since many psychologists must have thought about the problem at one time or another, this probably reflects the general belief that meanings are too complicated or too unique, or both. The few timid steps that have been taken in this direction involved drastic limitations on the scope of measurement, being aimed at scaling one or two isolated dimensions of meaning rather than meaning-in-general.

One group interested in scaling meaning has been the researchers in human learning, who wanted to be able to select materials for their experiments which could be specified with respect to this variable. A number of studies have been reported on the meaningfulness or "association value" of nonsense syllables (Glaze, 35; Hull, 42; Witmer, 113). The typical method was to use nonsense syllables as stimuli for word-associations, "meaningfulness" being indexed by proportions of subjects who could find any associations. One could then select equated lists for learning experiments, equated on this one basis at least.

Also motivated to provide learning experimenters with standardized

materials, Haagen (37) scaled 400 pairs of common adjectives in terms of their synonymity, vividness, familiarity, and association value. The method used was to have 280 college undergraduates judge these words on defined scales: (a) *synonymity* of a given word was judged on a seven-point scale in terms of the degree to which it denoted the same actions, objects or conditions as a standard word; (b) *vividness*, also judged on a seven-point scale, was defined as the clarity of graphicness of the impressions which a given word aroused; (c) *familiarity* was judged on a five-point scale, defined as the degree to which the judge knew the meaning of the word; and (d) *association value*, judged on a seven-point scale, referred to the degree to which the given word and a standard were associated in thought (e.g., hungry-thirsty, big-large, would have high association value). Useful though synonymity and associative value may be for purposes of learning experiments, they do not offer anything in the way of a measure of meaning—these judgments were always relative to some particular standard word, varying from one set of test words to another. The familiarity measure has nothing to do with meaning, of course. The vividness scale, being applied to each word separately rather than comparatively, probably is tapping some generalizable dimension of meaning.

Mosier (79) made the most direct application of scaling methods to the study of meaning. College subjects rated some 296 adjectives on an 11-point scale in terms of their favorableness-unfavorableness, these adjectives being selected from Thorndike's word lists as words expressing some degree of general evaluation. Frequency distributions of the responses to each word were scaled according to the method of successive intervals and plotted on probability paper. Plots for approximately 200 of these words were linear, indicating normal distribution of the data, when treated in this manner. Most of the words showed the "precipice effect" at one side or the other of the midpoint of the scale, indicating a higher degree of agreement on the *direction* (favorable or unfavorable) of the evaluation than on the *intensity*. Mosier was able to demonstrate a reasonable ordering of evaluative words in terms of their mean locations (e.g., excellent, good, common, fair, poor, etc.), including such information as the fact that "better" is connotatively less favorable than "good" (grammarians to the contrary). The most significant point is that Mosier demonstrated the feasibility of scaling certain aspects of meaning.

Summary on Existing Methods

The purpose of the preceding review has been to see if there already exist adequate methods of measuring meaning. By "adequate" I mean

already meeting most of the criteria of satisfactory measuring instruments. What are these criteria? (a) *Objectivity*. The method should yield quantitative and verifiable (reproducible) data. (b) *Reliability*. It should yield the same values within acceptable margins of error, when the same conditions are duplicated. (c) *Validity*. The data obtained should be demonstrably covariant with those obtained with some other, independent index of meaning. (d) *Sensitivity*. The method should yield differentiations commensurate with the natural units of the material studied, i.e., should be able to reflect as fine distinctions in meaning as are typically made in communicating. (e) *Comparability*. The method should be applicable to a wide range of phenomena in the field, making possible comparisons among different individuals and groups, among different concepts, and so on. (f) *Utility*. It should yield information relevant to contemporary theoretical and practical issues in an efficient manner, i.e., it should not be so cumbersome and laborious as to prohibit collection of data at a reasonable rate. While this is not an exhaustive listing of criteria, it is sufficient for our purposes.

1. The *physiological measures* (including action potential; GSR, and salivary records) are of somewhat dubious validity, since there has been no demonstration of the necessity of these peripheral components, and they are not sensitive measures in that we are unable to interpret details of the records in our present ignorance. Their chief drawback, however, is cumbersomeness—the subject has to be “rigged up” in considerable gadgetry to make such measurements. For this reason, even should validity and sensitivity problems be met satisfactorily, it seems likely that physiological indices will be mainly useful as criteria against which to evaluate more practicable techniques.

2. *Learning measures* (including semantic generalization and transfer/interference methods) are also somewhat cumbersome procedurally, but their main drawback as general measures of meaning is their lack of comparability. Any measure of generalization or interference is made with respect to the original learning of some standard which necessarily varies from case to case. The chief usefulness of learning measures, therefore, lies in the test of specific hypotheses.

3. The chief drawback with *perception measures* (e.g., what is perceived in ambiguous stimulus forms, the recognition-times for tachistoscopically presented words) is that they are not valid measures of *meaning*. They get at the availability or comparative habit strengths of alternate meanings or ways of perceiving. The fact that a religious person perceives VESPERS with a shorter presentation time than a theoretically oriented person says nothing about *how* the meaning of this term differs for them; the fact that the religious person perceives VESPERS more quickly than THEORY says nothing about the difference in meaning of these two words to this individual. The same statements apply to Skinner's “verbal summator” technique.

4. The selection of responses in *association methods* is partly dependent upon the meaning of the stimulus items (and hence indexes meaning) and partly dependent upon habit strength factors. The chief drawback, as a general measure of meaning, is lack of comparability. The responses of two individuals to the same stimulus, or of the same individual to two stimulus words, are

essentially unique as bits of data. Comparability can be obtained with group data, but this limits the method.

5. *Scaling methods* can be viewed as forms of controlled association in which the nature of the association is specified by definition of the scales (favorable-unfavorable, vividness, etc.) but the direction and intensity of association is unspecified. By the very nature of the scaling method, the comparability criterion is usually satisfied (provided the subjects can be shown to agree upon the meaning of the scale and its divisions). As used by Mosier, however, the method can have only partial validity. This is because he tapped only one dimension of meaning, the admittedly important evaluative dimension, whereas we know that meanings vary multidimensionally.

THE SEMANTIC DIFFERENTIAL

The method to be proposed here is a combination of associational and scaling procedures. It is an indirect method in the same sense that an intelligence test, while providing objective and useful information, does not directly measure this capacity. However, unlike the intelligence test which treats this ability *as if* it were distributed along a single continuum (e.g., IQ scores vary along a single scale), we accept at the outset that meanings vary in some unknown number of dimensions and frame our methodology accordingly.

Research Origins of the Method

This method had its origins in research on synesthesia, defined by Warren in his *Dictionary of Psychology* (109) as "a phenomenon characterizing the experiences of certain individuals, in which certain sensations belonging to one sense or mode attach to certain sensations of another group and appear regularly whenever a stimulus of the latter type occurs." This implies a sort of "neural short-circuiting" that is present in only a few freak individuals, and it is true that many of the classic case histories in this area gave credence to this view: a subject reported pressure sensations about his teeth and cheeks whenever cold spots on his arms were stimulated (Dallenbach, 27); a girl displayed a rigid system of relations between specific notes on the musical scale and specific color experiences, consistent when tested over a period of seven and one-half years (Langfeld, 60). But here, on the other hand, was a man who imagined the number "1" to be yellow, "2" to be blue, "3" to be red . . . and, of course, "8" to be black (anyone who has played pool will recognize the origin of this system); and here was a little girl who recalled her friends as having pink faces and her enemies as having purple faces. What modalities are crossed in these cases?

A more recent series of investigations by Karwoski, Odbert, and their associates related synesthesia to thinking and language in general (cf., also Wheeler and Cutsforth, 115). Rather than being a rare

phenomenon, Karwoski and Odbert (52) report that as many as 13 per cent of Dartmouth College students regularly indulged in color-music synesthesia, often as a means of enriching their enjoyment of music. These photistic visualizers varied among themselves as to the modes of translation employed and the vividness of their experiences, and their difference from the general population appeared to be one of degree rather than kind. Whereas fast, exciting music might be pictured by the synesthete as sharply etched, bright red forms, his less imaginative brethren would merely agree that terms like "red-hot," "bright," and "fiery," as verbal metaphors, adequately described the music; a slow and melancholic selection might be visualized as heavy, slow-moving "blobs" of sombre hue and described verbally as "heavy," "blue," and "dark." The relation of this phenomenon to ordinary verbal metaphor is evident: a happy man is said to feel "high," a sad man feels "low"; the pianist travels "up" and "down" the scale from treble to bass; souls travel "up" to the good place and "down" to the bad place; hope is "white" and despair is "black." The process of metaphor in language as well as in color-music synesthesia can be described as the parallel alignment of two or more dimensions of experience, defined verbally by pairs of polar opposites, with translations occurring between equivalent portions of the continua (Karwoski, Odbert, and Osgood, 53, pp. 212-221).

Interrelationships among color, mood, and musical experiences were studied more analytically by Odbert, Karwoski, and Eckerson (81). Subjects first listened to 10 short excerpts from classical scores and indicated their dominant moods by checking descriptive adjectives arranged in a mood circle (cf., Hevner, 39). Then, on a second hearing, they listed the colors appropriate to each score. Significant relations were shown; the color associations to musical scores followed the moods created. A portion of Delius' *On Hearing the First Cuckoo in Spring* was judged leisurely in mood and preponderantly green in color; a portion of Wagner's *Rienzi Overture* was judged exciting or vigorous in mood and preponderantly red in color. When another group of subjects was merely shown the mood adjectives (with no musical stimulation) and asked to select appropriate colors, even *more* consistent relations appeared, suggesting that the unique characteristics of the musical selections had, if anything, somewhat obscured the purely verbal or metaphorical relations between colors and moods. Almost identical findings have been reported by Ross (92) for relationships between the colors used in stage lighting and reported moods produced in the audience. Data are also available for the effects of color upon mood in mental institutions and in industrial plants.

Responses to complex selections of music such as used in the above studies are themselves too complex for analysis of specific relations between auditory-mood variables and color-form variables. In order to get closer to the mechanisms of translation, Karwoski, Odbert, and Osgood (53) used simple melodic lines recorded by a single instrument (clarinet) as stimuli. In a first experiment the subjects were typical photistic visualizers and they drew their photisms with colored pencils after hearing each short selection in a darkened room. The simplest

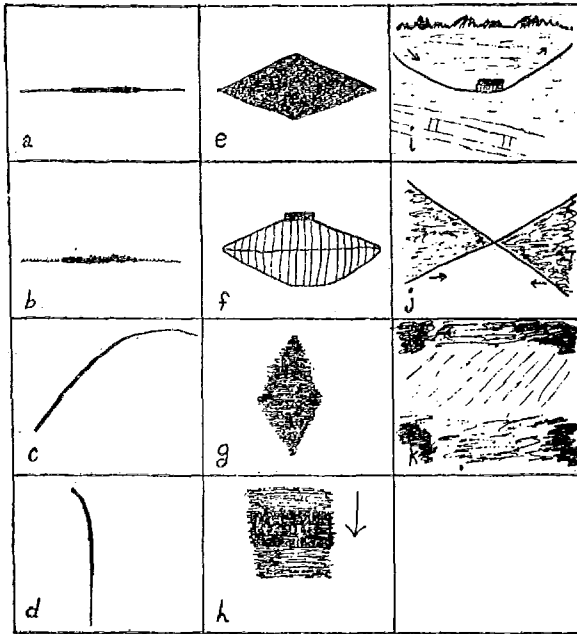


FIG. 2. SAMPLE OF PHOTISMS DRAWN BY COMPLEX SYNESTHETES TO REPRESENT A SIMPLE TONE WHICH GROWS LOUDER AND THEN SOFTER.

stimulus was a combination of crescendo and diminuendo on a single note—the sound merely grew louder, then softer—and this will serve to illustrate the results. As shown in Figure 2, subject *a* indicates increasing LOUDNESS by making the center of his line heavier, subject *b* by increasing amplitude of vibration, subjects *e*, *f*, and *g* by greater thickness of a solid form, subject *j* by more concentrated focusing, and subject *h* by more saturated coloring of the central portion. Subject *i* always created meaningful rather than abstract forms—here, a little car that comes nearer and then away again—yet the formal characteristics of his productions were generally like those of abstract synesthetes.

These are functionally or meaningfully equivalent responses to the same auditory stimulus dimension—i.e., there are alternate visual continua that can be paralleled with the loud-soft auditory continuum—and the advantage of the method is that its simplicity allows these relationships to show up clearly.

Are these photistic visualizers exercising a "rare" capacity or are they merely expressing overtly modes of translation that are implicit in the language of our culture? A second experiment used subjects who had never even thought of "seeing things" when they heard music (if they reported any such tendencies, they were eliminated). The same simple melodic lines as above were played and the subjects were instructed to "force themselves to draw something to represent what they heard." They produce the same types of visual forms and in approximately the same relative frequencies as the experienced visualizers. Finally, a group of 100 unselected students was given a purely verbal *meaning-polarity test*, each item of which appeared in the following form: LARGE-small; SOFT-LOUD, with instructions to circle that word in the second pair which "seems most clearly related to" the capitalized word in the first pair. Here again, essentially the same relations between music-mood variables and color-form variables discovered among sensitive synesthetes were linked meaningfully on the polarity test. *Large* was linked to *loud* by 96 per cent of these subjects, *near* with *fast* by 86 per cent, *bright* with *happy* by 96 per cent, *treble* with *up* by 98 per cent, and so on. It seems clear from these studies that the imagery found in synesthesia is on a continuum with metaphor, and that both represent *semantic* relations.

Are such semantic relations entirely dependent upon culture or is it possible that they reflect more fundamental determinants common to the human species? In an attempt to get at this question, the writer studied anthropological field reports on five quite widely separated primitive cultures—Aztec and Pueblo Indian, Australian Bushman, Siberian Aborigine, Negro (Uganda Protectorate), and Malayan—with the view of obtaining evidence on semantic parallelism. Special emphasis was given to nonmaterial aspects of culture (mythology, religion, arts, medical beliefs, birth, marriage, death complexes, etc.). The numerous pitfalls in the way of such analysis are probably obvious. Particularly, there is the danger of attributing relations to a primitive group when they are actually projections on the part of the observer or borrowings from the dominant Western culture. Therefore the results should be considered merely suggestive.

Nevertheless, the generality of certain relationships was quite striking. For example, *good* gods, places, social positions, etc., were

regularly *up* and *light* (*white*) in relation to *bad* things, which were *down* and *dark* (*black*). A prevalent myth tells of how the gods helped the original man to struggle "up" from the "dark," "cold," "wet," "sad" world below the ground to the "light," "warm," "dry," "happy" world on the surface of the earth. Among certain Siberian Aborigines, members of a privileged clan call themselves the "white" bones in contrast to all others who are referred to as "black" bones. And even among the Uganda Negroes we find some evidence for a white god at the apex of the hierarchy, and white cloth is clearly associated with purity, being used to ward off evil spirits and disease. Such data suggest the existence of a pervasive semantic frame of reference. Further study of the problem by more adequately trained investigators could be richly rewarding.

Stagner and Osgood (100) adapted this method and the logic underlying it to the study of social stereotypes. The notion of a continuum between the polar terms was made explicit by using such terms to define the ends of 7-step scales. Rather than studying the relations between continua, as above, a set of scales was used to measure the "meaning" of particular concepts, such as PACIFIST, RUSSIAN, DICTATOR, and NEUTRALITY. Successive samples of subjects were tested between April, 1940, and March, 1942 (including a sample obtained just prior to the Pearl Harbor incident). A single item on the tests appeared as follows:

PACIFIST: Kind : : : : : : cruel

with the subject instructed to check that position on the scale which best represented the direction and intensity of his judgment. The concepts and scales related in successive items of the test were randomized to insure as much independence of judgment as possible. The feasibility and efficiency of using this method to record the changing structures of social stereotypes (e.g., the changing meanings of a set of social signs) were demonstrated. That a total shift from an essentially pacifistic to an essentially militaristic frame of reference had been accomplished, even before the Pearl Harbor incident provided the spark to overt expression, was clearly evident in the data.

More important from the point of view of methodology was the following observation: As used by our subjects in making their judgments, the various descriptive scales fell into highly intercorrelated clusters. Fair-unfair, high-low, kind-cruel, valuable-worthless, Christian-anti-Christian, and honest-dishonest were all found to correlate together .90 or better. This cluster represented, we assumed, a single, general factor in social judgments, the evaluative (good-bad) dimension of the

frame of reference. Gradients like strong-weak, realistic-unrealistic, and happy-sad were independent of this evaluative group and pointed to the existence of other dimensions within the semantic framework. Enforced shifts in the apparent reference point of the observer (by having subjects judge the same concepts "as a German" or "as an Englishman") produced gross and appropriate changes in the evaluative dimension but did not disrupt the qualitative pattern of each stereotype—e.g., the stereotype GERMANS, when judged by students playing the role of Germans, was still seen as relatively more "strong" and "happy" (remember, this was during 1940–1942) than "noble" or "kind." This illustrates the kind of difficulty experienced when one tries to assume the point of view of another (cf., Stagner and Osgood, 99).

Logic of the Proposed Method

The researches described above gave rise to the following hypotheses:

1. *The process of description or judgment can be conceived as the allocation of a concept to an experiential continuum, definable by a pair of polar terms.* An underlying notion in our research is that these "experiential continua" will turn out to be reflections (in language) of the sensory differentiations made possible by the human nervous system. In other words, it is assumed that discriminations in meaning, which is itself a state of awareness, cannot be any finer or involve any more variables than are made possible by the sensory nervous system (cf., Boring, *The Dimensions of Consciousness*, 6). While failure to confirm this notion would not eliminate the proposed method as an index of meaning, its confirmation would greatly enhance the theoretical implications of this work.

2. *Many different experiential continua, or ways in which meanings vary, are essentially equivalent and hence may be represented by a single dimension.* This functional equivalence of many alternate continua was clearly evident in both the studies on synesthesia and those on the changing structure of social stereotypes. It is this fact about language and thinking that makes the development of a quantitative measuring instrument feasible. If the plethora of descriptive terms we utilize were in truth unique and independent of one another, as most philosophers of meaning seem to have assumed, then measurement would be impossible.

3. *A limited number of such continua can be used to define a semantic space within which the meaning of any concept can be specified.* From the viewpoint of experimental semantics, this both opens the possibility of measuring meaning-in-general objectively and specifies factor analysis as the basic methodology. If it can be demonstrated that a limited number of dimensions or factors are sufficient to differentiate among the meanings of randomly selected concepts, and if the technique devised satisfies the criteria of measurement stated earlier, then such a "semantic differential," as I have termed it, is an objective index of meaning. From the viewpoint of psychological theory, we may look upon the procedures followed in obtaining this measure as an operational definition of meaning, in the same sense that the procedures followed in obtaining the IQ score provide an operational definition of intelligence.

The operations followed in the present instance are explicit. They involve the subject's allocation of a concept within a standard system of descriptive dimensions by means of a series of independent associative judgments. The judgmental situation is designed to be maximally simple. Presented with a pair of descriptive polar terms (e.g., *rough-smooth*) and a concept (e.g., LADY), the subject merely indicates the direction of this association (e.g., LADY-*smooth*). We have developed two different methods for collecting data: In the *graphic method*, a pencil-and-paper technique which has the advantage that data can be collected from groups of subjects and hence very speedily, the subject indicates the intensity of his association by the extremeness of his checking on a 7-step scale. In the *judgment-time method*, which has the advantage that the subject cannot anticipate what concept is to be judged on a particular scale and hence cannot rationalize his reaction, intensity of association is indicated by the latency of the individual subject's choice reaction toward one or the other of the polar terms. In both methods each associative judgment of a particular concept against a particular descriptive scale constitutes one item. In successive items, concepts and dimensions are paired in deliberately rotated orders until every concept has been associated with every scale by every subject.

A Factor Analysis of Meaning

The procedures and results of this factor analysis will be described in detail elsewhere. A total of 50 descriptive scales, selected in terms of their frequency of usage, have been used in the judgment of 20 varied concepts, yielding a 1,000-item test. One hundred college students served as subjects. The graphic method was used.⁶ The purpose of this factor analysis is to isolate a limited number of general dimensions of meaning having a maximal differentiating power, to try to bring some order out of semantic chaos. The larger the proportion of total variance in meaning accounted for by these factors, the more satisfactory will be the measuring instrument finally set up. A preliminary estimation of factors in the 50×50 matrix (each scale correlated with every other scale) indicates the existence of several roughly independent dimensions. An "evaluative factor" accounts for by far the largest portion of the variance. There is also evidence for a "strength factor," an "activity factor," and several others not clearly defined in this rough approxima-

⁶ Apparatus for obtaining latency measurements from individual subjects has been constructed and will be standardized upon the reduced set of descriptive scales we hope to derive from this preliminary factor analysis. While this apparatus has the advantage that materials are projected from a film-strip and responses (directions and latencies) are photographed by a single-frame camera—all automatically—it is still applicable only to a single subject at a time and hence is time-consuming.

tion. Given such factors, it will be possible to select those specific scales (e.g., good-bad, strong-weak, active-passive, smooth-rough, hot-cold, etc.) which best represent them.

We have done some exploratory work on the use of the semantic differential as a practical measuring device. The two sets of profiles in Figure 3 will serve to illustrate the method. Two groups of only 20

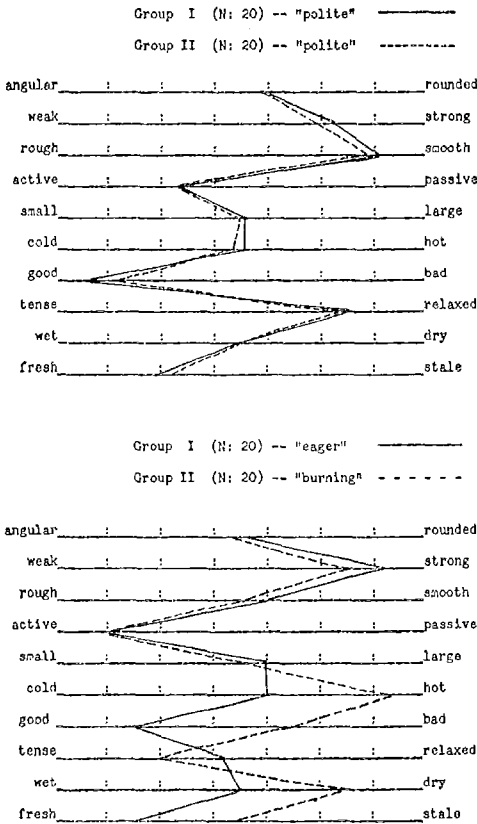


FIG. 3. ILLUSTRATION OF APPLICATION OF A PRELIMINARY FORM OF THE SEMANTIC DIFFERENTIAL FOR MEASURING THE CONNOTATIVE MEANINGS OF ADJECTIVES: A. UPPER PROFILES, MEDIANS FOR TWO GROUPS OF 20 SUBJECTS DIFFERENTIATING "POLITE"; B. LOWER PROFILES, MEDIANS FOR SAME TWO GROUPS OF 20 SUBJECTS DIFFERENTIATING "EAGER" AND "BURNING."

subjects each differentiated the meaning of the adjective "polite"; Group I also differentiated "eager" while Group II also differentiated "burning" (as part of a larger study). Median judgments of the 20 subjects on each scale are plotted. There is high agreement on the dif-

ferentiation of the same sign, "polite." The different, but somewhat similar signs, "eager" and "burning," show significant points of discrimination: whereas they are equally *strong* and *active*, "burning" is relatively *hot* and *dry* as compared with "eager," and "eager" is relatively *good* and *fresh* as compared with "burning." These differences are obviously what they would have to be if the method has any validity. It must be emphasized that the sample of scales shown here does *not* necessarily represent those to be finally derived from our factor analysis.

Evaluation of the Method

Evaluation of this instrument against the criteria of measurement listed earlier will be the subject of future reaserch, but some evidence can be presented now.

1. *Objectivity.* The semantic differential yields quantitative data which are presumably verifiable, in the sense that other investigators can apply the same sets of scales to equivalent subjects and obtain essentially the same result.

2. *Reliability.* In the test form from which data for the factor analysis were collected, 40 of the 1,000 items were selected at random and repeated. None of the subjects was aware that this had been done. The reliability coefficient was .85. The minimum variation in profile for the two groups of 20 subjects judging "polite" is another indication of the stability of the method.

3. *Validity.* All of the data collected so far on several problems display convincing face-validity and several direct experimental checks are planned. These include (a) correlation of attitudes toward various social objects as measured on standard tests with allocation of signs of these social objects within the semantic differential, and (b) the use of experimentally induced changes in meaning of signs (cf., Stagner and Britton, 98). We are not concerned about the problem of "labeling" factors, a point where the precision gained by the factor analytic method is often lost in the obscurities of language. Selection of specific scales to match factors can proceed on a purely objective basis, in terms of the factor loadings for each scale. As a matter of fact, the polar terms which define the scales do not admit much in the way of misinterpretation.

4. The question of *sensitivity* of the method comes down to whether it is able to reflect as fine distinctions in meaning as are ordinarily made. We have incidental evidence that a semantic differential can tease out nuances in meaning which are clearly felt but hard to verbalize deliberately.⁶ If there is a real difference in the meaning of two signs, such that they would not be used in

⁶ By way of illustration, most English-speaking Americans feel that there is a difference, somehow, between "good" and "nice" but find it difficult to explain. We gave several people these words to differentiate and it turned out that wherever "male" and "female" show a significant divergence, there also were "good" and "nice" differentiated (e.g., "good," like "male," is somewhat stronger, rougher, more angular, and larger than is "nice," which like "female" shifts toward the weak, smooth, rounded, and small directions of the space). Thus "nice man" has a slightly effeminate tone whereas "good woman" (as compared with "nice woman") has a narrowly moral tone.

precisely the same contexts, and if our measuring instrument includes a sufficient number of dimensions of the semantic space, then a significant difference should appear on at least one of the scales.

5. *Comparability.* It is here that the most serious questions arise. (a) *Is the method culture-bound?* If the tendency to dichotomize experiential continua is characteristic of Western culture but not necessarily elsewhere, then the method would not have generality. This is an empirical question requiring the skills of anthropologists and linguists for solution. (b) *Is the method limited to the differentiation of nouns against adjective scales?* The structure of our language is such that "adjectives" typically reflect abstracted qualities of experience and "nouns" the concepts and things dealt with. We have found it possible to set up scales like giant-midget, fire-iceberg, god-devil and to judge "concepts like INSINUATE and AGITATED against them. This does not seem "natural" to members of our language community, however; it is probably the stem or root meaning of words that our method taps. (c) *Can different concepts be compared?* To the extent that judgments of different concepts involve the same factor structure, any concept may be compared with any other against a single, standardized semantic framework. (d) *Can different individuals be compared?* This also comes down to the generality of the semantic factor structure. It is quite conceivable that different classes of people (scientists, ministers, etc.) have somewhat varied semantic structures, differing in the emphasis upon certain factors and interrelationships among them. In fact, a significant source of individual differences may lie here.

Our method can be criticized on the ground that it only gets at *connotative* meaning, not *denotative* meaning. This is a limitation. Both SIMON LEGREE and WAR might be allocated to approximately the same point in semantic space by our method. This would indicate similar connotative meaning, to be sure, but it would not indicate that these signs refer to the same object. Our differential will draw out the *hard, heavy, cold, ugly, threatening* connotations of the sign HAMMER, but it will not indicate that HAMMER is "an instrument for driving nails, beating metals, and the like, consisting of a head, usually of steel, fixed crosswise to a handle" (Webster's *Collegiate Dictionary*). In part, this limitation stems from our method of selecting descriptive scales in terms of frequency of usage rather than in terms of a logically exhaustive coverage, as given in Roget's *Thesaurus*, for example.

6. *Utility.* In any area of science, the development of an adequate method of measuring something (be it the wave length of radiation, blood chemistry, intelligence, or meaning) opens up well-nigh inexhaustible possibilities for application. (a) *Semantic norms.* In much the same way that Thorndike has established his norms for frequency-of-usage of common words in the English language, the semantic differential could be used to compile a functional lexicon of connotative meanings, a quantized thesaurus. Similarly, the gradual drift of changing meanings, both temporally and geographically, could be charted. (b) *Individual differences in meaning.* It is a truism that the meanings of socially significant signs differ for different classes of people. Concepts like CHURCH, LABOR LEADER, STALIN, and TRUMAN have different connotative significance to different people, and the semantic differential can be used to quantify these differences. In this sense, it can be used as a generalized, multidimensional attitude test. For example, 10 people may have identical degrees of favorableness toward NEGRO (evaluative dimension) and yet vary markedly with respect to other dimensions of the meaning-space. (c) *Changes*

in meaning. Under the pressure of events, the meanings of social signs change, e.g., the meaning of ITALIANS to Americans during the past half century. Similarly, under the "pressure" of psychotherapy, the meaning or emotional significance of certain critical concepts (e.g., FATHER, THERAPIST, ME, etc.) undergoes change. (d) *Quantification of subjective language data.* We have recently used the semantic differential as a means of scoring TAT reactions; not only is the testing process greatly speeded up, but the data are in easily manipulable form. Preliminary studies indicate that the essential individual differences in meaning of such projective materials, as teased out of complicated verbatim "stories," are sharply etched in the semantic differential data. (e) *Cross-cultural communication problems.* If the structure of the semantic space proves to be sufficiently general that the method can be translated into equivalent differentials in other languages, numerous possibilities are opened up. Are the fundamental factors in meaning and their relationships independent of the language spoken? Can the significant points of deviation in meaning of critical concepts, as between Americans and Russians, for example, be discovered? Can the finer, subtler degrees of acculturation into a new society be traced? And there are other potential applications, to aesthetics, to studying the development of meaning in children, and so on.

SUMMARY

The first portion of this paper describes a behavioral conception of the sign-process as developed from a general mediation theory of learning. The remainder is concerned with the problem of measuring meaning. Various existing approaches to the problem—physiological, learning, perception, association, and scaling methods—have been evaluated against the usual criteria of measurement and have been found inadequate. The development of a semantic differential as a general method of measuring meaning is described. It involves (a) the use of factor analysis to determine the number and nature of factors entering into semantic description and judgment, and (b) the selection of a set of specific scales corresponding to these factors which can be standardized as a measure of meaning. Using this differential, the meaning of a particular concept to a particular individual can be specified quantitatively as a particular point in the multidimensional space defined by the instrument. Some of the possible uses of such a measuring instrument are briefly indicated.

BIBLIOGRAPHY

1. ALLPORT, G. W. *Personality: a psychological interpretation.* New York: Holt, 1937.
2. BARTLETT, F. C. *Remembering.* Cambridge, England: Cambridge Univer. Press, 1932.
3. BINGHAM, W. E., JR. A study of the relations which the galvanic skin response and sensory reference bear to judgments of the meaningfulness, significance, and importance of 72 words. *J. Psychol.*, 1943, 16, 21-34.
4. BLOOMFIELD, L. *Language.* New York: Holt, 1933.
5. BODER, D. P. The adjective-verb-

- quotient: a contribution to the psychology of language. *Psychol. Rec.*, 1940, 3, 310-343.
6. BORING, E. G. *The physical dimensions of consciousness*. New York: Appleton-Century, 1933.
 7. BOUSFIELD, W. A. An empirical study of the production of affectively toned items. *J. gen. Psychol.*, 1944, 30, 205-215.
 8. BOUSFIELD, W. A. The relationship between mood and the production of affectively toned associates. *J. gen. Psychol.*, 1950, 42, 67-85.
 9. BOUSFIELD, W. A., & BARRY, H., JR. Quantitative correlates of euphoria. *J. exp. Psychol.*, 1937, 21, 218-222.
 10. BOUSFIELD, W. A., & SEDGWICK, C. H. W. An analysis of sequences of restricted associative responses. *J. gen. Psychol.*, 1944, 30, 149-165.
 11. BOUSFIELD, W. A., & BARCLAY, W. D. The relationship between order and frequency of occurrence of restricted associative responses. *J. exp. Psychol.*, 1950, 40, 643-647.
 12. BRUNER, J. S., & GOODMAN, C. C. Value and need as organizing factors in perception. *J. abnorm. soc. Psychol.*, 1947, 42, 33-44.
 13. BRUNER, J. S., & POSTMAN, L. Symbolic value as an organizing factor in perception. *J. soc. Psychol.*, 1948, 27, 203-208.
 14. BRUNSWIK, E. Die Zugänglichkeit von Gegenständen für die Wahrnehmung. *Arch. ges. Psychol.*, 1933, 88, 377-418.
 15. CARMICHAEL, L., HOGAN, H. P., & WALTER, A. A. An experimental study of the effect of language on the reproduction of visual perceived form. *J. exp. Psychol.*, 1932, 15, 73-86.
 16. CARROLL, J. B. Diversity of vocabulary and the harmonic series law of word-frequency distribution. *Psychol. Rec.*, 1938, 2, 379-386.
 17. CARROLL, J. B. The analysis of verbal behavior. *Psychol. Rev.*, 1944, 51, 102-119.
 18. CARTER, L. F., & SCHOOLER, K. Value, need and other factors in perception. *Psychol. Rev.*, 1949, 56, 200-207.
 19. CASON, H., & CASON, E. B. Association tendencies and learning ability. *J. exp. Psychol.*, 1925, 8, 167-189.
 20. CHOTLOS, J. W. Studies in language behavior: IV. A statistical and comparative analysis of individual written language samples. *Psychol. Monogr.*, 1944, 56, No. 2 (Whole No. 255), 75-111.
 21. CLARK, H. M. Conscious attitudes. *Amer. J. Psychol.*, 1911, 22, 214-249.
 22. CLARK, R. S. An experimental study of silent thinking. *Arch. Psychol.*, 1922, 7, No. 48. Pp. 101.
 23. COFER, C. N., & FOLEY, J. P., JR. Mediated generalization and the interpretation of verbal behavior. I. Prolegomena. *Psychol. Rev.*, 1942, 49, 513-540.
 24. COFER, C. N., JANIS, M. G., & ROWELL, M. M. Mediated generalization and the interpretation of verbal behavior. III. Experimental study of antonym gradients. *J. exp. Psychol.*, 1943, 32, 266-269.
 25. COMSTOCK, C. On the relevancy of imagery to the processes of thought. *Amer. J. Psychol.*, 1921, 32, 196-230.
 26. CROSLAND, H. R. A qualitative analysis of the process of forgetting. *Psychol. Monogr.*, 1921, 29, No. 1 (Whole No. 130).
 27. DALENBACH, K. M. Synaesthesia: "Pressury" cold. *Amer. J. Psychol.*, 1926, 37, 571-577.
 28. DARROW, C. W., & HEATH, L. L. Reaction tendencies relating to personality. In C. P. Stone, C. W. Darrow, C. Landis, and L. L. Heath. *Studies in the dynamics of*

- behavior. Chicago: Univer. Chicago Press, 1932.
29. DORCUS, R. M. Habitual word associations to colors as a possible factor in advertising. *J. appl. Psychol.*, 1932, 16, 277-287.
 30. DUNCKER, K. On problem-solving. (Trans. by L. S. Lees from 1935 original.) *Psychol. Monogr.*, 1945 58, No. 5 (Whole No. 270).
 31. Estes, W. K. A visual form of the verbal summator. *Psychol. Rec.*, 1940, 4, 174-180.
 32. FOLEY, J. P., JR., & COFER, C. N. Mediated generalization and the interpretation of verbal behavior: II. Experimental study of certain homophone and synonym gradients. *J. exp. Psychol.*, 1943, 32, 169-175.
 33. FOLEY, J. P., JR., & MACMILLAN, Z. L. Mediated generalization and the interpretation of verbal behavior: V. "Free association" as related to differences in professional training. *J. exp. Psychol.*, 1943, 33, 299-310.
 34. FOLEY, J. P., JR., & MATHEWS, M. Mediated generalization and the interpretation of verbal behavior: IV. Experimental study of the development of interlinguistic synonym gradients. *J. exp. Psychol.*, 1943, 33, 188-200.
 35. GLAZE, J. A. The association value of non-sense syllables. *J. genet. Psychol.*, 1928, 35, 255-269.
 36. GOODWIN, J., LONG, L., & WELCH, L. Generalization in memory. *J. exp. Psychol.*, 1945, 35, 71-75.
 37. HAAGEN, C. H. Synonymity, vividness, familiarity, and association value ratings of 400 pairs of common adjectives. *J. Psychol.*, 1949, 27, 453-463.
 38. HEBB, D. O. *The organization of behavior*. New York: Wiley, 1949.
 39. HEVNER, KATE. Experimental studies of the elements of expression in music. *Amer. J. Psychol.*, 1936, 48, 246-268.
 40. HOWES, D. H., & SOLOMON, R. L. A note on McGinnies' "Emotional-ity and perceptual defense." *Psychol. Rev.*, 1950, 57, 229-234.
 41. HOWES, D. H., & SOLOMON, R. L. Visual duration threshold as a function of word-probability. *J. exp. Psychol.*, 1951, 41, 401-410.
 42. HULL, C. L. The meaningfulness of 320 selected nonsense syllables. *Amer. J. Psychol.*, 1933, 45, 730-734.
 43. HULL, C. L. *Principles of behavior: An introduction to behavior theory*. New York: Appleton-Century, 1943.
 44. JACOBSON, E. Electrophysiology of mental activities. *Amer. J. Psychol.*, 1932, 44, 677-694.
 45. JOHNSON, W. Studies in language behavior: I. A program of research. *Psychol. Monogr.*, 1944, 56, No. 2 (Whole No. 255), 1-15.
 46. JONES, H. E. Emotional factors in learning. *J. gen. Psychol.*, 1929, 2, 263-272.
 47. JONES, H. E., & WECHSLER, D. Galvanometric technique in studies of association. *Amer. J. Psychol.*, 1928, 40, 607-612.
 48. JUNG, C. G. *Studies in word-association*. (Trans. by M. D. Eder.) London: William Heinemann, 1918.
 49. KAPUSTNIK, O. P. The interrelation between direct conditioned stimuli and their verbal symbols (trans. from Russian title), 1930. *Psychol. Abstracts*, 1934, 8, No. 153.
 50. KARWOSKI, T. F., & BERTHOLD, F., JR. Psychological studies in semantics: II. Reliability of free association tests. *J. soc. Psychol.*, 1945, 22, 87-102.
 51. KARWOSKI, T. F., GRAMLICH, F. W., & ARNOTT, P. Psychological studies in semantics: I. Free association reactions to words, drawings, and

- objects. *J. soc. Psychol.*, 1944, 20, 233-247.
52. KARWOSKI, T. F., & ODBERT, H. S. Color-music. *Psychol. Monogr.*, 1938, 50, No. 2 (Whole No. 222).
53. KARWOSKI, T. F., ODBERT, H. S., & OSGOOD, C. E. Studies in synthetic thinking: II. The roles of form in visual responses to music. *J. gen. Psychol.*, 1942, 26, 199-222.
54. KARWOSKI, T. F., & SCHACHTER, J. Psychological studies in semantics: III. Reaction times for similarity and difference. *J. soc. Psychol.*, 1948, 28, 103-120.
55. KELLER, MARGARET. Mediated generalization: the generalization of a conditioned galvanic skin response established to a pictured object. *Amer. J. Psychol.*, 1943, 56, 438-448.
56. KENT, GRACE H., & ROSANOFF, A. J. A study of association in insanity. *Amer. J. Insanity*, 1910, 67, 37-96, 317-390.
57. KOFFKA, K. *Principles of gestalt psychology*. New York: Harcourt Brace, 1935.
58. KÖHLER, W. *The mentality of apes*. (Trans. by E. Winter.) New York: Harcourt Brace, 1925.
59. KOTLIAREVSKY, L. I. Cardio-vascular conditioned reflexes to direct and to verbal stimuli (trans. from Russian title), 1936. *Psychol. Abstracts*, 1939, 13, No. 4046.
60. Langfeld, H. S. Note on a case of chromaesthesia. *Psychol. Bull.*, 1914, 11, 113-114.
61. Lazarus, R. S., & McCLEARY, R. A. Autonomic discrimination without awareness: A study of subception. *Psychol. Rev.*, 1951, 58, 113-122.
62. LEVINE, R., CHEIN, I., & MURPHY, G. The relation of the intensity of a need to the amount of perceptual distortion: a preliminary report. *J. Psychol.*, 1942, 13, 283-293.
63. LYNCH, C. A. The memory values of certain alleged emotionally toned words. *J. exper. Psychol.*, 1932, 15, 298-315.
64. McCLELLAND, D. C., & ATKINSON, J. W. The projective expression of needs: I. The effect of different intensities of the hunger drive on perception. *J. Psychol.*, 1948, 25, 205-222.
65. MCGINNIES, E. Emotionality and perceptual defense. *Psychol. Rev.*, 1949, 56, 244-251.
66. MCGINNIES, E. Discussion of Howes' and Solomon's note on "Emotionality and perceptual defense." *Psychol. Rev.*, 1950, 57, 235-240.
67. MAIER, N. R. F. Reasoning in humans. III. The mechanisms of equivalent stimuli and of reasoning. *J. exp. Psychol.*, 1945, 35, 349-360.
68. MALINOWSKI, B. The problem of learning in primitive languages. Supplement in C. K. Ogden, & I. A. Richards. *The meaning of meaning*. New York: Harcourt Brace, 1938.
69. MARBE, K. *Experimentell-psychologische Untersuchungen über das Urteil*. Leipzig: Engelmann, 1901.
70. MASON, M. Changes in the galvanic skin response accompanying reports of changes in meaning during oral repetition. *J. gen. Psychol.*, 1941, 25, 353-401.
71. MAUSNER, B., & SIEGEL, A. The effect of variation in "value" on perceptual thresholds. *J. abnorm. soc. Psychol.*, 1950, 45, 760-763.
72. MAX, L. W. An experimental study of the motor theory of consciousness. III. Action-current responses in deaf-mutes during sleep, sensory stimulation, and dreams. *J. comp. Psychol.*, 1935, 19, 469-486.
73. MAX, L. W. An experimental study of the motor theory of consciousness. IV. Action-current responses in the deaf during awakening,

- kinaesthetic imagery, and abstract thinking. *J. comp. Psychol.*, 1937, **24**, 301-344.
74. MELTON, A. W., & IRWIN, J. MCQ. The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific response. *Amer. J. Psychol.*, 1940, **53**, 173-203.
75. METZNER, C. A. The influence of preliminary stimulation upon human eyelid responses during conditioning and during subsequent heteromodal generalization. *Summ. Doct. Diss. Univ. Wis.*, 1942, **7**, 152-154.
76. MILLER, G. A. *Language and communication*. New York: McGraw-Hill, 1951.
77. MORRIS, C. *Foundations of the theory of signs*. Internat. Encycl. Unif. Sci., Vol. 1, No. 2. Chicago: University of Chicago Press, 1938.
78. MORRIS, C. *Signs, language and behavior*. New York: Prentice-Hall, 1946.
79. MOSIER, C. I. A psychometric study of meaning. *J. soc. Psychol.*, 1941, **13**, 123-140.
80. MOWRER, O. H. Preparatory set (expectancy): Some methods of measurement. *Psychol. Monogr.*, 1940, **52**, No. 2 (Whole No. 233).
81. ODBERT, H. S., KARWOSKI, T. F., & ECKERSON, A. B. Studies in synesthetic thinking: I. Musical and verbal associations of color and mood. *J. gen. Psychol.*, 1942, **26**, 153-173.
82. OGDEN, C. K., & RICHARDS, I. A. *The meaning of meaning*. London: Kegan Paul, 1923.
83. OSGOOD, C. E. Meaningful similarity and interference in learning. *J. exp. Psychol.*, 1946, **36**, 277-301.
84. OSGOOD, C. E. An investigation into the causes of retroactive interference. *J. exp. Psychol.*, 1948, **38**, 132-154.
85. POSTMAN, L., & BRUNER, J. S. Perception under stress. *Psychol. Rev.*, 1948, **55**, 314-324.
86. POSTMAN, L., BRUNER, J. S., & MCGINNIES, E. Personal values as selective factors in perception. *J. abnorm. soc. Psychol.*, 1948, **43**, 142-154.
87. RAZRAN, G. H. S. Salivating and thinking in different languages. *J. Psychol.*, 1935-1936, **1**, 145-151.
88. RAZRAN, G. H. S. A quantitative study of meaning by a conditioned salivary technique (semantic conditioning). *Science*, 1939, **90**, 89-90.
89. REED, H. B. The existence and function of inner speech in thought processes. *J. exp. Psychol.*, 1916, **1**, 365-392.
90. RIESS, B. F. Semantic conditioning involving the galvanic skin reflex. *J. exp. Psychol.*, 1940, **26**, 238-240.
91. RIESS, B. F. Genetic changes in semantic conditioning. *J. exp. Psychol.*, 1946, **36**, 143-152.
92. ROSS, R. T. Studies in the psychology of the theater. *Psychol. Rec.*, 1938, **2**, 127-190.
93. SANFORD, R. N. The effects of abstinence from food upon imaginal processes: A preliminary experiment. *J. Psychol.*, 1936, **2**, 129-136.
94. SANFORD, R. N. The effects of abstinence from food upon imaginal processes: A further experiment. *J. Psychol.*, 1937, **3**, 145-159.
95. SKINNER, B. F. The verbal sunimator and a method for the study of latent speech. *J. Psychol.*, 1936, **2**, 71-107.
96. SKINNER, B. F. The distribution of associated words. *Psychol. Rec.*, 1937, **1**, 69-76.
97. SKINNER, B. F. *The behavior of organisms*. New York: Appleton-Century-Crofts, 1938.
98. STAGNER, R., & BRITTON, R. H., JR. The conditioning technique applied to a public opinion problem. *J. soc. Psychol.*, 1949, **29**, 103-111.

99. STAGNER, R., & OSGOOD, C. E. An experimental analysis of a nationalistic frame of reference. *J. soc. Psychol.*, 1941, 14, 389-401.
100. STAGNER, R., & OSGOOD, C. E. Impact of war on a nationalistic frame of reference: I. Changes in general approval and qualitative patterning of certain stereotypes. *J. soc. Psychol.*, 1946, 24, 187-215.
101. SWITZER, S. C. A. Disinhibition of the conditioned galvanic skin response. *J. gen. Psychol.*, 1933, 9, 77-100.
102. THORNDIKE, E. L. *The teacher's word book*. New York: Teachers College, Columbia Univ., 1921.
103. THORNDIKE, E. L., & LORGE, I. *The teacher's word book of 30,000 words*. New York: Bureau of Publications, Teachers College, Columbia Univ., 1944.
104. THORSON, A. M. The relation of tongue movements to internal speech. *J. exp. Psychol.*, 1925, 8, 1-32.
105. THOULESS, R. H. Phenomenal regression to the real object. II. *Brit. J. Psychol.*, 1931, 22, 1-30.
106. THUMB, A., & MARBE, K. *Experimentelle Untersuchungen über die psychologischen Grundlagen der sprachlichen Analogiebildung*. Leipzig: W. Engelmann, 1901.
107. TRAUGOTT, N. N. The interrelations of immediate and symbolic projections in the process of the formation of conditioned inhibition (trans. from Russian title), 1934. *Psychol. Abstracts*, 1935, 9, No. 1166.
108. TRAUGOTT, N. N., & FADEYEVA, V. K. The effect of difficult extinction of food-procuring conditioned reflexes upon the general and speech behavior of children (trans. from Russian title), 1934. *Psychol. Abstracts*, 1935, 9, No. 1167.
109. WARREN, H. C. (Ed.) *Dictionary of psychology*. New York: Houghton Mifflin, 1934.
110. WASHBURN, M. *Movement and mental imagery*. Boston: Houghton Mifflin, 1916.
111. WATSON, J. B. *Behavior, an introduction to comparative psychology*. New York: Holt, 1914.
112. WERTHEIMER, M. *Productive thinking*. New York: Harper, 1945.
113. WITMER, L. R. The association value of three-place consonant syllables. *J. gen. Psychol.*, 1935, 47, 337-360.
114. WOODWORTH, R. S. *Experimental psychology*. New York: Holt, 1938.
115. WHEELER, R. H., & CUTSFORTH, T. D. Synaesthesia and meaning. *Amer. J. Psychol.*, 1922, 33, 361-384.
116. WYCZOIKOWSKI, A. Theoretical and experimental studies in the mechanism of speech. *Psychol. Rev.*, 1913, 20, 448-458.
117. ZIFF, G. K. *The psychobiology of language*. New York: Houghton Mifflin, 1935.
118. ZIFF, G. K. *Human behavior and the principle of least effort*. Cambridge, Mass.: Addison-Wesley, 1949.

Received August 29, 1951.