INTRODUCTION

If a psychologist is asked how a “perception” can be observed, he will usually reply that we really observe a verbal report or some other overt response taking place in the presence of certain stimulus patterns. Such an acknowledgment is what often passes for an “operational definition” of perception. Actually such apparent scientific sophistication has not always been significant within the psychology of perception. It is so easy to say that one could make such an operational definition if one really wanted to, that much of the literature on perception remains without specified operational definition of concepts. Perceptual processes like “seeing,” “distorting,” and “selecting,” are concepts which are still used incorrectly in the psychology of perception. Such concepts often are private and imprecise, although presumably they need not be; and they often admit ad hoc explanations of certain observed relations between responses and stimuli. In the recent literature on perception, for example, we find statements such as the following: “A perception is an experience of something” (2, p. 14). “What one sees, what one observes, is inevitably what one selects from a near infinitude of potential percepts” (9, p. 142). “The goal of perception, in its broadest sense, is the construction of a meaningful behavioral environment—an environment congruent with ‘reality’ on the one hand and the needs and dispositions of the organism on the other . . .” (2, p. 314). Such statements have been used as definitions of concepts in the theoretical formulation of certain experimental observations. Our feeling is that, if the operations that are basic to these concepts had been specified carefully, these statements would not have been made. They would have been recognized as redundancies, or at best, as assertions for which evidence is lacking in our present stage of knowledge.

We believe it is particularly important that the observed properties of those psychological events that are often described in such a quasi-subjective manner be specified with great care. For instance, if one takes the operational definition of concepts very seriously, it will be clear that a “perception” is defined by specifiable stimulus properties and specifiable response properties. One familiar type of experiment within the field of perception is characterized by the fact that human subjects say or do something in the presence of a situation established partly by optical operations and partly by verbal-instructional ones. Thus one is driven to specify the response properties that define perceptual concepts as well as the situational properties. It is all too conventional at present for perceptual behavior to be interpreted exclusively in terms of the situational properties—and, unhappily, for the situational properties to be defined in terms of private phenomenology.

1 The second author is now at Tulane University.
2 This research was supported by the Laboratory of Social Relations, Harvard University. The authors are indebted to Professor Leo Postman for many helpful suggestions in the preparation of this paper.
In this paper we are interested in the properties of a particular class of responses that can be labelled linguistic, since a predominance of recent perceptual data on human beings has been derived from the relationship between spoken reports and optical situations. Thus, the study of perception becomes in this case a study of the properties of linguistic responses relative to varying stimulus configurations. This paper is an attempt to show that so-called "perceptual phenomena" can sometimes be deduced directly from the relationship between certain known properties of the linguistic responses of individuals and certain properties of the optical and instructional situation in which those responses occurred. From the above analysis of the operations which define perceptual concepts it follows that in perceptual experiments in which linguistic responses are elicited from human subjects, any variable that is a general property of linguistic responses must also be a property of any perceptual concept that is based upon those responses. Here we shall consider one linguistic variable in particular: The frequency of occurrence of a word in a general sample of the English language. The perceptual concept based upon linguistic operations that we shall consider is the duration for which the printed form of a word must be exposed in a tachistoscope, under a given set of instructions, before a subject will emit that word. This is called the visual duration threshold of that word. We have treated the relation between these two variables elsewhere (4). In the present paper we wish to extend the analysis developed in that paper to a somewhat more complicated topic in the field of perception.

To be more specific, let us examine a recent paper by Postman, Bruner, and McGinnies (9). These experimenters concluded that an individual's "interests" or "values" affect the duration thresholds for recognition of words that represent those values. They defined the duration threshold as the tachistoscopic exposure time required by a subject to recognize an exposed word. From the obtained relationship between values and visual duration thresholds these authors concluded that there were two perceptual processes in operation: perceptual selectivity and perceptual defense. Perceptual selectivity refers to the fact that in their experiment relatively short mean duration thresholds were found for words which represented those areas of interest that a subject valued highly. Perceptual defense refers to the fact that they found that words of low value areas had relatively long mean duration thresholds for correct report.

Postman, Bruner, and McGinnies went through the following experimental procedures. First, the Allport-Vernon Study of Values (1) was given to their subjects. This questionnaire investigates a subject's interests relative to six general areas: theoretical, economic, aesthetic, social, political, and religious. On the basis of scores obtained on this test, the orientation of each subject toward these six areas of interest was ranked from one to six.4

4 To clarify the terminology in this type of measurement, we shall use the word interest to designate the six fields of orientation; the word value will be reserved for the relative degree of orientation of a subject toward any one of these six areas of interest as indicated by his score on the Allport-Vernon questionnaire. Thus one can speak of the value a subject attributes to a particular area of interest, or one can speak of an individual whose two highest values lie in the theoretical and
The experimenters then presented to their subjects 36 fairly common English words (six words representing each of the six interest areas), exposing these words in a tachistoscope and using an ascending method-of-limits procedure. The duration thresholds for each of the 36 words were obtained by determining the exposure duration which was just necessary to enable the subject to report correctly the word that was exposed in the tachistoscope. A slight tendency was found for words representing highly valued interests to have lower duration thresholds than words in lowly valued areas. A graph of this result is given in Fig. 1. To account for their findings the authors introduced two perceptual mechanisms. (1) "Value orientation acts as a sensitizer, lowering thresholds for acceptable stimulus objects" (e.g., words in a highly valued area) (9, p. 151). They named this mechanism selective sensitization. (2) "Value orientation may, on the other hand, raise thresholds for unacceptable stimulus objects" (9, p. 151). This mechanism was called perceptual defense. Both of these hypothesized perceptual mechanisms must operate under subthreshold stimulation, since they obviously cannot be operative after the subject has met the threshold criterion by reporting a stimulus word correctly.

It must be emphasized that Postman, Bruner, and McGinnies derived their concepts from data consisting of linguistic responses to linguistic stimuli of a visual nature. Is it possible to account for their findings in terms of known properties of this relationship? We believe it is. It has long been known that the familiarity of an object is inversely related to the speed with which that object can be recognized. More specifically, we have shown in another paper (4) that the frequencies of occurrence of words in the Thorndike-Lorge word counts (11) show an inverse relation to the visual duration thresholds for those words. Sample data from this study are shown in Fig. 2.

Now let us assume that high valuation of a given area of interest is associated with a positive deviation from the mean frequencies with which words

---

**Fig. 1.** Visual duration threshold (in seconds) as a function of Allport-Vernon value rank. These data are from Postman, Bruner, and McGinnies (6).

**Fig. 2.** Visual duration threshold (in seconds) as a function of Thorndike-Lorge word frequency. These data are from Howes and Solomon (3).
in that area occur in general usage. This assumption simply means that people who are highly interested in a subject generally use words associated with that subject more often than people who have no special interest in the subject. Positive and negative deviations in word-frequency thus would characterize the words that represent the extreme value ranks one and six of the Allport-Vernon test. The Allport-Vernon test is especially appropriate to this assumption, since it was validated on professional groups who were selected to represent the various categories of interest because they spent their full time in activities concerned with those interests.

This assumption, coupled with the experimental finding that an increase in the frequency of words is accompanied by a decrease in their duration thresholds, permits us to deduce the results of the experiment by Postman, Bruner, and McGinnies without having to postulate the two mechanisms of perceptual selectivity and perceptual defense. High value rank is correlated with a relatively high positive deviation from general word-usage—i.e., with relatively frequent use of a word. Frequency of use of a word is inversely related to the word's visual duration threshold, hence, the inverse relation between visual duration threshold and value rank.

This deduction of the Postman, Bruner, and McGinnies result has the advantage that it leads to certain other theorems that are susceptible to experimental investigation. From (a) the experimental relationship referred to above between visual duration thresholds for words and the Thorndike-Lorge word frequencies, and (b) the assumption that a given increase in the value placed upon a certain area of interest is correlated with a corresponding increase in word frequency for all words in that area, it follows that:

(1) The difference between the duration thresholds for words representing two areas of interest will be greater when there are relatively large differences between the actual value scores obtained for those areas of interest on the Allport-Vernon questionnaire. Thus the relation between duration threshold and value rank will depend upon the test scores from which the value ranks are derived. Subjects who show extreme scores on the Allport-Vernon test should yield larger differences between the mean duration thresholds for value ranks one and six, for example, than subjects who show only small variations in their value scores.

Since the relationship between visual duration threshold and word frequency approximates an inverse logarithmic function (see 4), a given change in word frequency will produce a smaller change in duration threshold if it occurs for a word of relatively high frequency than if it occurs for a word of relatively low frequency. This consideration leads to a second deduction:

(2) A given difference between two value ranks will produce a smaller difference in duration thresholds for relatively frequent words representing the two areas of interest than in the thresholds for relatively infrequent words representing those interests. A quantitative treatment of this deduction is presented in Appendix 1.

The fact that in our experiments we were able to obtain stable relationships between visual duration threshold and word frequency with less than 20 subjects is strong evidence that the frequency of occurrence of a word varies only slightly from person to person in comparison to extreme variations of frequency of occurrence from word to word. This conclusion also has a strong basis in everyday observation of language behavior. If we accept this inference, we can state a third deduction:

(3) Differences between the duration
thresholds of words representing extreme differences in value rank will be small compared with differences between the thresholds of words representing extreme differences in word frequency in the Thorndike-Lorge word counts.

Now let us turn to the experimental test of these three deductions.

**PROCEDURE**

This experiment is essentially a repetition of the experiment reported by Postman, Bruner, and McGinnies (9). Their stimulus words, however, were uncontrolled for frequency of occurrence in the Thorndike-Lorge word counts. Consequently we devised a new list of words representing the six areas of interest. To be certain that the words we used could be called representative of the fields of interest to which they were assigned, we selected words that were used to represent the respective interests in the actual questions on the Allport-Vernon test. Five words were so chosen for each of the six interest areas. These 30 words constitute the relatively frequent words of our experiment, and they are listed in Table I. Each of the 30 relatively infrequent words shown in Table I was chosen to match one of the 30 frequent ones with respect to interest category, and to differ from it in having a considerably lower frequency of occurrence in the Thorndike-Lorge counts. The procedure was to select a relatively infrequent word from the lists of cognate words found in a standard thesaurus (8). Synonyms were chosen wherever possible. Care was taken that each relatively infrequent word did not differ greatly in length from its relatively frequent cognate. No words less than six letters long were selected.

At each experimental session the subject was first given the Allport-Vernon Study of Values. Then the 60 words listed in Table I were presented in the tachistoscopic procedure. This procedure has been described in detail in a previous paper (4). In brief, each word was first exposed for a duration too short to permit the subject to report it correctly. The duration of succeeding exposures was then increased gradually until the subject reported the

<table>
<thead>
<tr>
<th>Table I</th>
<th>Stimulus Words</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Theoretical</td>
</tr>
<tr>
<td>Relatively frequent</td>
<td>scientific</td>
</tr>
<tr>
<td></td>
<td>physics</td>
</tr>
<tr>
<td></td>
<td>intellectual</td>
</tr>
<tr>
<td></td>
<td>knowledge</td>
</tr>
<tr>
<td></td>
<td>education</td>
</tr>
</tbody>
</table>

(Mean log frequency = 2.23)

| Relatively infrequent | inductive | limousine | uncoerced | elegies | chancells |
| | statics | broker | barrister | vignette | celestial |
| | percipience | frugality | rebuttal | etcher | psychical |
| | erudition | assets | assidious | ensemble | beatific |
| | pedagogue | mundane | judiciary | metaphor | theistic |
| | | | | | condolence |

(Mean log frequency = 0.60)
word correctly. A duration threshold for each word was obtained separately in this way for each subject. Nineteen subjects were used.

RESULTS

The value ranks for each subject were derived from his scores obtained on the six areas of interest in the Allport-Vernon test. The highest value area for a given subject was given a rank of one and the lowest a rank of six. The visual duration thresholds for words representing each value rank were then averaged over the group of 19 subjects, regardless of what specific field of interest happened to receive that value rank. For example, if Subject A scored highest on the Allport-Vernon test in the area of theoretical interest and Subject B scored highest in the aesthetic area of interest, then Subject A's duration thresholds for the theoretical words (scientific, physics, intellectual, knowledge, education) would be averaged with Subject B's thresholds for the aesthetic words (poetry, picture, painter, orchestra, literary) in order to obtain the mean duration threshold for frequent words in value rank one. Thresholds for frequent and for infrequent words were averaged separately. Thus the mean duration threshold for each value rank, either for frequent or for infrequent words, is based upon a total of 95 threshold measurements. These means, along with the means for frequent and infrequent words combined (n = 190) are shown in Table II.

Inspection of these data shows no indication of systematic variation of visual duration threshold with value rank. There is, however, an appreciable difference between the mean thresholds for frequent and infrequent words at every value rank. The results shown in Table II thus substantiate our third prediction, that threshold differences associated with differences in value rank will be small compared with those associated with differences in word frequency. It is clear that the data in

<table>
<thead>
<tr>
<th></th>
<th>Value ranks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Frequent words</td>
<td>.132</td>
</tr>
<tr>
<td>Infrequent words</td>
<td>.211</td>
</tr>
<tr>
<td>Means</td>
<td>.172</td>
</tr>
</tbody>
</table>

\[\text{TABLE II}\]

Mean Duration Thresholds (in Seconds) for Frequent and Infrequent Words Stimuli of Each Value Rank
Table III

Mean Duration Thresholds (in Seconds) for Frequent and Infrequent-Word Stimuli in Each Value Rank for a Selected Sample of Eleven Subjects

<table>
<thead>
<tr>
<th>Value rank</th>
<th>(High) 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>(Low) 6</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequent words</td>
<td>.114</td>
<td>.105</td>
<td>.119</td>
<td>.112</td>
<td>.131</td>
<td>.125</td>
<td>.118</td>
</tr>
<tr>
<td>Infrequent words</td>
<td>.195</td>
<td>.206</td>
<td>.182</td>
<td>.211</td>
<td>.204</td>
<td>.221</td>
<td>.203</td>
</tr>
</tbody>
</table>

Table II do not bear out those reported by Postman, Bruner, and McGinnies which were plotted in Fig. 1. However, what indication there is of a general tendency in the data seems to point in the direction of lower thresholds for the high value ranks: The mean duration threshold for the highest three value ranks is slightly lower than the mean duration threshold for the lowest three ranks both for the frequent and the infrequent words. The mean duration thresholds for value ranks one, two, and three combined, for the frequent and the infrequent words, are 0.128 and 0.207 second, respectively. For value ranks four, five, and six combined, the corresponding means are slightly higher, 0.131 and 0.218 second, respectively.

Now if our first deduction is substantiated, it should be possible to enlarge the threshold differences from value rank to value rank by selecting the data for those subjects who show extreme scores on the Allport-Vernon test. We therefore set aside for separate treatment those eleven subjects whose Allport-Vernon scores exhibited unusually large departures in any interest area from the standardized norms for that test (see 1). The duration thresholds for these eleven selected subjects are given in Table III and are plotted against value rank in Fig. 3. An examination of Fig. 3 shows that the trend for the eleven selected subjects is obviously in the direction found by Postman, Bruner, and McGinnies. When

Table IV

Analysis of Variance of Data in Table III

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of squares</th>
<th>df.</th>
<th>Mean square</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>129,728.39</td>
<td>654</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>S's</td>
<td>38,278.77</td>
<td>10</td>
<td>3,827.88</td>
<td>32.53</td>
<td>*</td>
</tr>
<tr>
<td>F</td>
<td>12,278.77</td>
<td>1</td>
<td>12,883.26</td>
<td>109.49</td>
<td>*</td>
</tr>
<tr>
<td>V's</td>
<td>1,138.29</td>
<td>5</td>
<td>227.66</td>
<td>1.93</td>
<td>—</td>
</tr>
<tr>
<td>S x F</td>
<td>6,208.92</td>
<td>10</td>
<td>620.92</td>
<td>5.28</td>
<td>*</td>
</tr>
<tr>
<td>S x V</td>
<td>2,845.61</td>
<td>50</td>
<td>56.91</td>
<td>.48</td>
<td>—</td>
</tr>
<tr>
<td>F x V</td>
<td>946.21</td>
<td>5</td>
<td>189.24</td>
<td>1.61</td>
<td>—</td>
</tr>
<tr>
<td>Error</td>
<td>67,427.33</td>
<td>573</td>
<td>117.67</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

* Significant at 1 per cent level.
The three highest and three lowest value ranks are compared, the differences between mean duration thresholds for high and low value ranks are greater than those observed for all 19 subjects. The mean duration thresholds for value ranks one, two, and three combined, for the frequent and the infrequent words, are 0.113 and 0.194 second, respectively. For value ranks four, five, and six combined, the corresponding means are 0.123 and 0.212 second, respectively. The fact that the relation between duration threshold and value rank is clearer with subjects selected for extreme value scores is in agreement with the first deduction from the assumptions discussed in the introduction.

The data on the 11 selected subjects are also congruent with our second deduction; since the infrequent words exhibit considerably greater variation in mean duration threshold as a function of value rank than do the frequent words. In addition, the third deduction receives strong support from these data, since they show that even when subjects are selected for their extreme variations in Allport-Vernon score the differences in duration thresholds associated with differences in value rank are quite small compared with the difference between duration thresholds for frequent and infrequent words. Thus all three deductions correspond with the actual tendencies observed in our experimental data.

An analysis of variance for the data of the eleven selected subjects is summarized in Table IV. By far the largest estimate of variance is that associated with differences in word frequency or familiarity. Individual differences among subjects also yielded a large estimate of variance. The F-ratios for both of these sources of variance are significant at the 0.01 level. The differences among duration thresholds for words in the different value ranks are not significant according to the F-test. However, as we have pointed out, there is a marked trend for high-value words to have shorter duration thresholds. The difference between mean duration thresholds for infrequent words of value ranks one and six is significant at the 0.05 level.

The extremely low variance attributable to the interaction between subjects and value ranks indicates that the subjects behaved quite uniformly with respect to the way their thresholds vary

**Fig. 3.** Visual duration threshold (in seconds) as a function of Allport-Vernon value rank with word frequency as a parameter. The upper curve (hollow circles) is for relatively infrequent words (mean log word frequency = 0.60). The lower curve (black dots) is for relatively frequent words (mean log word frequency = 2.23). The straight lines have been fitted by the method of least squares. They serve merely as convenient indications of central tendencies of the data, and should not be interpreted to signify a rectilinear function between duration threshold and value rank, since value rank is not a continuous dimension and the size of intervals between ranks is unspecified.

6 This is true whether the F-ratios are based upon the error term or upon the significant interaction variance attributable to subject X word frequency.
with value rank. The interaction between frequency and value rank, although it is not statistically significant, is large in comparison with the variance contributed by value rank alone. This result may indicate that the difference between the two slopes in Fig. 3 might well be statistically significant in an experiment in which value rank emerged as a variable significantly associated with duration thresholds.

Of our three deductions, only the third—concerning the relative importance of word frequency and value rank as variables related to variation in duration thresholds—is adequately tested by our statistical analysis. Its confirmation is striking. The present data do not permit an adequate statistical test of the first deduction: that value ranks based upon the extreme scores in the Allport-Vernon test will be associated with greater differences in duration thresholds than will value ranks based upon the relatively small differences in test scores. This deduction, however, is supported by a comparison of the data in Tables II and III. The second deduction, which requires that the slope of the relationship between duration thresholds and value ranks should vary with the Thorndike-Lorge frequencies of the two groups of words used in this experiment, cannot be tested legitimately by an analysis of variance for these data. This follows from the fact that we found no significant variation in value duration threshold to be associated with variation in value rank. The data of Table III, however, are certainly consistent with the deduction, and an examination of the quantitative data presented in Appendix I lends additional support to it.

Thus, in summary, the three deductions from our general assumptions describe the outstanding trends in the data, but additional experiments will certainly be necessary in order to determine the reliability of these tendencies for the first two deductions. It is also apparent from our results that the shift in duration threshold correlated with differences in personal values is quite small, and is probably more difficult to obtain than the earlier data of Postman, Bruner, and McGinnies would indicate.

DISCUSSION

1. The interpretation of word frequency. Experimental studies of the learning of linguistic materials have led to the use of two frequency variables: (a) the frequency of visual exposure of a linguistic stimulus, and (b) the frequency of occurrence of a linguistic response. It is theoretically possible to distinguish between these two frequency variables in any experiment utilizing linguistic materials, but a word count such as the Thorndike-Lorge list does not do this. For the sake of simplicity, we have in this paper considered Thorndike-Lorge word frequencies as estimates of the frequency of words “in a general sample of the English language.” This is, of course, an inexact formulation. The frequency of a word in such a list can be interpreted either as a stimulus or as a response variable. In the first case a Thorndike-Lorge frequency is considered as an estimate of the frequency with which a word is exposed during the reading of a population of subjects who read the class of printed material that is analyzed in the word count. In the second case the word-count frequency is considered as an estimate of the frequency with which a word is emitted by the population of subjects who wrote the material sampled by the count. It seems reasonable to expect that there would be a high correlation between these two frequency variables, but whether visual duration threshold is related to each of them in the same way remains to be seen. In the absence of further data it seems best
merely to define word frequency for English words as the frequency of words in the Thorndike-Lorge tables.

2. The Allport-Vernon test as a corollary of word frequency. "Value," "interest," and similar concepts, we have suggested, are often inferences from word frequencies. If a subject uses many economic words relative to words representative of other areas of interest, he is often classified as having economic "values." Scores on the Allport-Vernon test must somehow reflect such differences in word usage. In the introduction we considered this relation between Allport-Vernon score and word frequency as a separate assumption. In this section we wish to show that it can be stated as a corollary of differences in word frequency and the experimentally demonstrated function between duration threshold and word frequency. If this can be accomplished, only one principle will be necessary for the description of our own results and those of Bruner, Postman, and McGinnies.

Let us examine carefully the operations that determine what score a subject will receive when he is given the Allport-Vernon test. The Allport-Vernon test can be considered as a set of visual choice discriminations between alternative groups of words, the choice being presented as a series of paired comparisons. Each comparison is between words representing two different interest areas. For example, consider item 4 of the test: "If you were a university professor and had the necessary ability, would you prefer to teach: (a) poetry; (b) chemistry and physics?" If, in a series of items like this, a subject always chooses the poetry side of the question, we say that he has strong aesthetic values, and we assign a value rank of one to the aesthetic area for that subject. Value rank, therefore, reflects the frequency with which words representing one interest area are chosen in preference to words representing another interest area.

Now let us consider two populations of subjects, A and B, which are identical except for the fact that A uses the word poetry more frequently than B does, and that B uses chemistry and physics more often than A does. What will happen when they are presented with a questionnaire item such as item 4 from the Allport-Vernon test? We know from other data (4) that the duration threshold of poetry will be lower for A than for B, and that the duration thresholds for chemistry and physics will be lower for B than for A. Let us then imagine a questionnaire item in which the subject responds vocally to the printed form of the two possible choices instead of checking a number that corresponds to the printed form of the choice word (e.g., the numeral 1 to poetry). Since on this test the operations performed are identical to those defining duration threshold, we know from (4) that, following any given amount of exposure to the printed alternatives, the word with the higher frequency of use will tend to occur more often. For our two populations A and B, therefore, there will be a tendency for A to make the response "poetry" more often than B, and a tendency for B to make the response "chemistry and physics" more often than A. On a modified Allport-Vernon test composed of such items, therefore, population A will be said to have higher values in aesthetics than B, and B will be said to have higher values than A in the theoretical area, although it has been assumed that the only difference between the two populations is the frequency with which they use the two sets of words.

To generalize this argument from our hypothetical questionnaire to the actual Allport-Vernon test involves one major
assumption, viz., that the response of checking a number that corresponds to a printed word does not differ fundamentally from the response of emitting the vocal sounds; that it corresponds to the frequency with which that word has been emitted in the past as a "reading response" in the presence of the visual exposure of its written form. To make this interpretation it is necessary to postulate that covert ("implicit") word-responses, which cannot actually be observed, occur during reading. On the other hand, it permits value rank and duration threshold to be viewed as two derivatives of the strength of association between linguistic stimuli and covert linguistic responses. In that way it provides a unified description of the results presented above.

According to association theories, the more frequently a response has been associated in the past with a stimulus, the briefer the exposure that is necessary to elicit that response overtly. Now it is also possible to consider the Allport-Vernon questionnaire in the association paradigm. The Allport-Vernon scale can be considered as a set of visual choice discriminations between alternative groups of words, the choice being presented as a series of paired comparisons. For example, consider again item 4 of the test: "If you were a university professor and had the necessary ability, would you prefer to teach (a) poetry; (b) chemistry and physics?" The printed words poetry and chemistry or physics are the discriminative stimuli and the circling of an answer is the overt response in this discrimination. In conventional association interpretation this response is assumed to be mediated by the covert response of one of the alternative word-pairs, and the appropriate circling response will depend upon which mediating response is more strongly associated with its visual stimulation. According to all association theories of verbal learning the covert word-response that has been more often associated with its discrimination stimulus will tend to be chosen more often. We can as yet advance no experimental data for this assumption. Nevertheless both types of response are learned as appropriate responses to printed words, so that the association is of the same derivation in both cases. If we are willing to accept the assumption for the time being, we can deduce a subject's score on the Allport-Vernon test from the inverse relation between word frequency and duration threshold, plus a knowledge of how frequently that subject uses words representing the various areas of interest.

3. Antecedent conditions related to word frequency. Word frequency has been treated throughout as a property of the state of the organism at the time of measurement of duration thresholds. The interrelations of word frequency, duration threshold, and personal values are thus cross-sectional with respect to time. Frequency of occurrence in the Thorndike-Lorge counts, however, is not in itself an independent variable that is capable of experimental manipulation. Our analysis of the approximately simultaneous determinants of duration threshold, therefore, needs to be supplemented by methods of manipulating word frequency as it has been defined in the above experiments. Two general approaches may be suggested: (a) We can artificially establish differential word frequencies in a laboratory situation, observing their effects upon subsequent perceptual phenomena; (b) we can specify particular antecedent conditions (independent variables) that are sufficient to establish word frequencies as a variable dependent upon those conditions. When, after building word frequencies into the organism by these procedures, we can accurately predict related perceptual phenomena, we may
find that many so-called "determinants" of perception will be by-passed or will become redundant.

The first approach may be illustrated by a preliminary experiment (10). Word frequencies were produced artificially by having Turkish words—completely unfamiliar to English-speaking subjects—pronounced different numbers of times. Word frequency was in this way established as an independent variable capable of direct manipulation. Measurement of duration thresholds of those words followed, and the relation between the threshold for a word and its frequency of previous pronunciation was found to approximate the relation between duration thresholds for English words and Thorndike-Lorge word frequencies.

The second approach to the manipulation of word frequency depends upon its being related to a system of genetic or historical relationships, such as those basic to current learning theories. Among the best established propositions in modern association theory are those that relate the frequency of occurrence of a learned response (as a dependent variable) to specified conditions under which the response was learned (independent variables). These independent variables are extremely varied, including such factors as reinforcement, rate of repetition, time between repetitions, etc. A special case of this dependent variable of learning experiments is the frequency of occurrence of words, since words are learned responses. All of the independent variables above, consequently, should permit the control of word frequency and indirectly of duration thresholds for words if the propositions of learning theory concerning response probability are valid. When viewed from this perspective, the relation between word frequency and duration threshold is a case of the general relationship between the probability of a response and the latency of its occurrence in the presence of a discriminated stimulus. In this respect it bears upon a fundamental problem in learning theory, since response probability and response latency are sometimes considered as two manifestations of the single, basic concept of habit strength.

McGinnies (7) and others (6) have argued that "emotional" determinants operate selectively in the tachistoscopic situation in determining visual duration thresholds. We do not believe there is sufficient experimental evidence to justify such a conclusion. At present one can, with reasonable precision, account for the duration thresholds of words classified as emotional, as well as of neutral words, simply in terms of their Thorndike-Lorge frequencies, without making additional assumptions (5). The discussion of the preceding paragraph should make it clear that this does not mean that we consider "emotional" determinants unimportant to the analysis of language behavior. Emotional factors undoubtedly operate to an important extent in the building of word frequencies in a given life history. In this way they would be related to word frequency and, indirectly, to duration threshold in the manner suggested above for other antecedent conditions of learning. But to date we can find no evidence to suggest that emotional factors operate in the tachistoscopic situation independently of their effect upon word frequency.

The investigation of duration thresholds for words has only begun. We are well aware that the frequency of use of a word is not the only property of that word that is related to its duration threshold. Nor do we wish to over-emphasize the response attributes of perceptual data. Stimulus generalization, for instance, is also evident in tachistoscopic experiments such as the one we have described. Thus when the
word *surmise* is exposed very briefly in the tachistoscope, the word *surprise* is often the response that is given; the frequency with which such a generalized response occurs, moreover, appears to depend upon the Thorndike-Lorge frequency of that word, among other things. Many such properties of the events occurring in such experiments need to be investigated before an adequate theoretical account of the data can be given. Nonetheless, the relationships we have reported previously (4, 5), taken together with those presented in this paper, strongly suggest that the Thorndike-Lorge frequency of a word will probably turn out to be an important variable in any such theory. And the preceding discussion demonstrates how inquiring more deeply into the nature of the observations that determine certain customary psychological concepts may lead to a more coherent organization of the body of knowledge that concerns those concepts. From this point of view, it is reasonable to expect that the theoretical development of much of the psychology of human perception will be interrelated closely with the development of the psychology of language behavior.

**APPENDIX I**

Quantitative statement can be given to the second deduction, that the function relating duration threshold to value rank will vary with the Thorndike-Lorge frequency of the words representing the various fields of interest. First of all, let us restate the basic assumption made in the introduction to this paper, that, for words representing a field of interest, valuation of that interest is associated with a departure from the mean frequency of use in the general population. If $p$ is the probability (relative frequency) in the general population of any word in a given field of interest, a high value placed on that field by a subject thus will mean that the probability of such a word for that subject will be greater than $p$ by some quantity $\Delta p$. In terms of the operations employed in our experiment, a high score in theoretical interest on the Allport-Vernon test would mean that, on the average, the probabilities of words representing theoretical interest (e.g., scientific and inductive) will be increased over their Thorndike-Lorge values by some quantity $\Delta p$.

The second basic assumption concerns the relation between duration threshold and word frequency. This relation can be written (4)

$$t = -k \log p - C,$$  \hspace{1cm} (1)

in which $t$ is the mean duration threshold (in units of time), $f$ the Thorndike-Lorge frequency of the word, and $k$ and $C$ are constants. Considerable evidence indicates that this relationship holds over the entire range of word frequencies represented in the present experiment (3, 4). These data also show that the constant $C$ has the value zero when a constant, 4.4, is subtracted from $\log f$. Since this difference amounts to the division of a set of frequencies by a constant that is larger than any member of that set of frequencies, it assumes certain properties of a relative frequency or probability. Let $\rho$ stand for the probability of a word based on such a relative frequency. Then

$$\log \rho = \log f - 4.4$$ \hspace{1cm} (2)

and

$$t = -k \log \rho.$$ \hspace{1cm} (3)

Now let us take two words with probabilities $p_1$ and $p_2$ that represent the same area of interest. Our assumptions permit us to describe the probabilities of those words for an individual who has positive value in that field of interest as $(p_1 + \Delta p)$ and $(p_1 + \Delta p)$, respectively. The corresponding duration
thresholds are given by
\[ t_1 + \Delta t_1 = -k \log (p_1 - \Delta \phi), \]
\[ t_2 + \Delta t_2 = -k \log (p_2 - \Delta \phi), \]
in which \( t_1 \) and \( t_2 \) are the thresholds corresponding to \( p_1 \) and \( p_2 \), and \( \Delta t_1 \) and \( \Delta t_2 \) are the changes in threshold introduced by the addition of \( \Delta \phi \) to \( p_1 \) and \( p_2 \). Assigning the value zero to \( t_1 \) and \( t_2 \) to simplify comparison, we have
\[ \Delta t_1 = -k \log (p_1 + \Delta \phi), \]
\[ \Delta t_2 = -k \log (p_2 + \Delta \phi). \]
When \( \Delta \phi \) is small,
\[ \Delta t_1 / \Delta t_2 = \log p_1 / \log p_2. \]

To express this result in words, our assumptions imply this equation: The differences in duration thresholds for infrequent and for frequent words introduced by a given difference in value rank will stand in the same proportion to each other as the ratio of the logarithms of the Thorndike-Lorge probabilities of the two sets of words.

Mean logarithms of word frequencies for both frequent and infrequent words are shown in Table I. By Equation 2 the mean probability for the 30 infrequent words is \( 0.60 - 4.40 = -3.80 \); that for the 30 frequent words is \( 2.23 - 4.40 = -2.17 \). The ratio of these values of \( \log \rho \) is 1.75+. The difference in duration threshold associated with differences in value rank, \( \Delta t \), can best be computed from our data as the difference between the mean duration threshold for the highest three value ranks averaged together and that for the lowest three value ranks averaged together (\( N = 165 \) for each mean). Table V compares the ratio of log word probabilities with the corresponding ratio of differences in duration threshold. For the data on 11 subjects with extreme scores on the Allport-Vernon test, these ratios are almost identical. The threshold differences for the total group of 19 subjects are so small that they are less than the smallest unit of measurement to which duration thresholds were determined (0.01 sec.), but yield a ratio that is of the order of magnitude to be expected from the ratio of word probabilities.

Although these data are suggestive, they are of course insufficient to establish the particular set of assumptions we have employed. It is to be hoped

<table>
<thead>
<tr>
<th>Type of measurement</th>
<th>Duration thresholds</th>
<th>Ratio: Infrequent/Frequent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Ratios of log word probabilities</td>
<td>Infrequent</td>
<td>Frequent</td>
</tr>
<tr>
<td>2. Lowest 3 value ranks minus highest 3 (11 subjects)</td>
<td>0.018</td>
<td>0.010</td>
</tr>
<tr>
<td>3. Lowest 3 value ranks minus highest 3 (19 subjects)</td>
<td>0.011</td>
<td>0.003</td>
</tr>
<tr>
<td>4. Value rank 6 minus value rank 1 (least squares fit of Fig. 3)</td>
<td>0.027</td>
<td>0.015</td>
</tr>
</tbody>
</table>
that comparable data from other laboratories will soon permit a more extensive test of the deductions. Even at the present stage, however, we believe this formulation illustrates the potentiality of a concrete statement of the operations involved in the measurement of "perception." Such a statement makes it possible for a single unambiguous formulation to describe complex relationships among perceptual data which quasi-subjective formulations like perceptual defense and perceptual sensitization could not reveal.

REFERENCES


[MS. received September 19, 1950]