

The imagery effect and phenomenal background frequency in verbal discrimination learning

GEORGE D. GOEDEL and JUDITH A. ENGLERT
State University of New York, College at Geneseo, Geneseo, New York 14454

A series of four experiments was conducted to assess the role of phenomenal background frequency in verbal discrimination learning and its possible involvement in the imagery effect. The initial two experiments produced a reliable imagery effect for mixed and unmixed lists with respect to concreteness of pair members, regardless of phenomenal frequency manipulations, with words high in objective background frequency. No effects were found for phenomenal background frequency. Experiment 3 involved phenomenal frequency ratings for 200 abstract and 200 concrete words. Experiment 4 evaluated the role of phenomenal background frequency for a mixed list using words low in objective frequency. A reliable imagery effect was again found with no effects for phenomenal frequency. An alternative hypothesis involving differential accrual of situational frequency to abstract and concrete items during verbal discrimination learning to explain the imagery effect was also tested by Experiment 4 but was not supported by the data.

A rather consistent finding reported in the verbal discrimination learning (VDL) literature in recent years is the easier learning of concrete (or high-imagery) word pairs relative to abstract (or low-imagery) pairs in mixed or unmixed lists (Paivio & Rowe, 1970, 1971; Rowe & Paivio, 1971, 1972; Ullrich & Balogh, 1972). This finding has proven somewhat distressing for frequency theory (Ekstrand, Wallace, & Underwood, 1966; Wallace, 1972), which has thus far served well as the major theory of VDL, since the effect is neither predicted nor readily interpretable, in any obvious way, by the theory as it currently stands. The reliability of the effect would seem to demand a search for an interpretation that could be accommodated by frequency theory or a considerable revision of the theory if it is to survive without substantial limitations.

A recent report by Ghatala and Levin (1976) has suggested that the imagery effect may be due to an apparent confound of phenomenal background frequency with concrete and abstract items in such a way that the effect may be interpretable by appeal to the Weber's law postulate found in frequency theory. Should such be the case, the search for a viable frequency theory explanation would be ended and the theory would escape serious revision. Thus, a major purpose of the present investigation was a systematic assessment of this possibility.

We are grateful to R. G. Crowder and E. J. Rowe for their constructive criticisms of an earlier version of this manuscript and to Joel Levin for his encouragement and interest. We would also like to thank Christine Beard for collecting and partially analyzing the data in Experiment 4. Requests for reprints should be sent to the first author at the Department of Psychology, State University of New York, College at Geneseo, Geneseo, New York 14454. The second author is now at Southern Illinois University, Carbondale, Illinois.

The evidence directly indicating such a confound was provided by Galbraith and Underwood (1973). These investigators demonstrated that, although objective background frequency of concrete and abstract words was equated based on available word counts (Kučera & Francis, 1967; Thorndike & Lorge, 1944), college student subjects tended to perceive abstract words as having higher frequencies than concrete words. This phenomenon has since been confirmed in subsequent investigations (Ghatala & Levin, 1976; Goedel & Thomas, 1977) and indicates that "phenomenal" or perceived background frequency indeed differs from objective (word-count) frequency for this subject population, and that matching items on the latter alone may be an insufficient control for frequency differences.

Ghatala and Levin (1976) rightfully noted that all previous studies reporting the imagery effect in VDL failed to control for phenomenal frequency and that a confound was apparent. They also pointed out that since the Weber's law postulate of frequency theory predicts that ease of learning in a VDL task is inversely proportional to the base frequency of the items, a higher perceived preexperimental frequency for abstract words relative to that for concrete words could explain the so-called "imagery effect." It should perhaps be noted, however, that although the Weber's law postulate has received considerable support when item frequency is experimentally manipulated (situational frequency), the evidence suggesting an assimilation of background and situational frequency, and therefore, the operation of this postulate with respect to preexperimental or background frequency, has been rather weak and limited as given in extensive reviews by Eckert and Kanak (1974) and Wallace (1972). Nevertheless, recent evidence by Ghatala, Levin, and Makoid (1975) provides support for the Weber's law prediction with

preexperimental item frequency and thus makes reasonable an interpretation of the imagery effect in terms of the postulate.

Ghatala and Levin (1976) demonstrated that when only objective frequency of concrete and abstract words is controlled, allowing abstract items to have higher phenomenal background frequency, the imagery effect in VDL is obtained. However, when concrete and abstract words were matched on both objective and phenomenal background frequency, the usually reliable imagery effect did not reach an acceptable level of significance. They concluded that this evidence "provides strong support for the notion that phenomenal background frequency is an important variable in verbal discrimination learning" and that the imagery effect can be accounted for by the covariation of concreteness and phenomenal frequency.

The experiments performed in the present investigation were prompted by an unwillingness to accept the null effect obtained by Ghatala and Levin (1976) as strong evidence in support of their interpretation. It seemed more reasonable to us to directly manipulate phenomenal frequency in a factorial design with item concreteness in an attempt to demonstrate the importance of phenomenal frequency as a variable in VDL than to infer importance based on the lack of an imagery effect when this variable was controlled. In addition, we wished to evaluate whether or not phenomenal background frequency was independent of item concreteness or whether it interacted with this variable in some fashion to attenuate the imagery effect. Finally, since a number of VDL studies demonstrating the imagery effect have employed rather common words typically having A to AA Thorndike-Lorge frequency (e.g., Paivio & Rowe, 1971; Rowe & Paivio, 1971), a demonstration of a phenomenal frequency effect with such materials seemed warranted prior to attributing the imagery effect obtained in such studies to a phenomenal background frequency confound. Thus, the first two experiments of this investigation were performed to directly test the interpretation of the imagery effect offered by the Ghatala and Levin (1976) study using words with A to AA Thorndike-Lorge frequency. The third experiment involved the collection of phenomenal frequency ratings on a 5-point scale for 200 concrete and 200 abstract words in an attempt to replicate the Ghatala and Levin (1976, Experiment 1) findings and provide a pool of scaled materials for Experiment 4. The fourth and final experiment served two purposes. It provided a further exploration of the phenomenal frequency variable in a mixed-list design with words low in Thorndike-Lorge frequency, as well as an evaluation of an alternative hypothesis for the imagery effect, also consistent with frequency theory, suggesting a greater accrual of subjective frequency for abstract items relative to concrete during VDL.

EXPERIMENT 1

Method

Subjects. The subjects were 64 students at the State University of New York at Geneseo enrolled in introductory psychology classes. The subjects received course credit for their participation. None of the subjects had previous experience in verbal learning studies. They were assigned randomly to treatment conditions as they entered the laboratory, with the restriction that the $n+1$ th subject was not assigned until the n th subject had been tested in every condition.

Materials. Verbal materials consisted of 40 concrete (C) and 40 abstract (A) words drawn from the Spreen and Schulz (1966) norms, with mean ratings of 6.46 (range: 4.59-6.96) and 2.54 (range: 1.46-3.88), respectively, on a 7-point scale. All words had Thorndike-Lorge frequency values of A to AA and care was taken to match m values as closely as possible for the two classes of words. In addition, half of the C and A words were low in rated phenomenal frequency with means of 4.33 (range: 3.98-4.65) and 4.35 (range: 3.55-4.79), respectively, on a 9-point scale with the other half being high with means of 6.28 (range: 5.63-7.14) and 6.29 (range: 6.03-6.63), respectively, based on an earlier study (Goedel & Thomas, 1977) where phenomenal frequency data was gathered for C and A words.

Two lists of 20 word pairs, each homogeneous with respect to concreteness of pair members (CC or AA), were randomly constructed with the restriction that half of the pairs in each list contain words rated high in phenomenal frequency and half low. This procedure was repeated a second time to yield a replication of lists with different word pairings to control for idiosyncratic pairs. Thus, 16 subjects received each form of the two classes of word lists. Four study and four test orders of each list were constructed such that: (1) within each order, the occurrence of high and low phenomenal frequency pairs was random; (2) within an order, the arbitrarily selected correct items of each pair were located on the right for half of the high and low phenomenal frequency pairs and on the left for the other half; (3) for a random half of high and low phenomenal frequency pairs, the spatial location of the correct item changed between study and test orders; and (4) for all pairs, the spatial location of the correct item changed at least twice across study and test orders. Half of the subjects receiving each list condition received a second version of the list in which the other member of each pair was correct (i.e., arbitrarily selected correct and incorrect designations for word-pair members were reversed).

Procedure. Each subject was tested individually and received a total of four alternating study and test trials. The subject was seated across from the experimenter and in front of a BCI programmer (Model SR-400), which presented word pairs one at a time at a constant rate of 3 sec each. During study trials, the correct member of each pair was underlined and the subject was instructed to read each member of the pair aloud from left to right as it was presented and then to repeat aloud the underlined word. During test trials, each pair was presented without underlining and the subject was simply to respond aloud with the correct alternative. The sequence of study orders was varied across subjects and the words "study" and "test" preceded each study and test list for a 3-sec period. The experimenter recorded each response made by the subject on each test trial.

Design. This experiment essentially consisted of a 2 by 2 by 4 mixed factorial design. Word-pair concreteness (CC or AA) was manipulated between subjects, while phenomenal frequency of pair members (high or low) and VDL trials (four) served as within-subjects variables.

Results and Discussion

The data subjected to analysis of variance were the number of incorrect identifications (errors) made by

Table 1
Mean Total Errors as a Function of Pair Concreteness, Phenomenal Frequency, and Trials

| Pair Type | Trial 1 | | Trial 2 | | Trial 3 | | Trial 4 | | Mean |
|-----------|---------|------|---------|------|---------|------|---------|------|------|
| | High | Low | High | Low | High | Low | High | Low | |
| CC | 2.13 | 2.31 | 1.16 | 1.50 | 1.09 | 1.13 | .78 | 1.09 | 1.40 |
| AA | 2.72 | 2.69 | 2.38 | 2.16 | 2.59 | 2.06 | 2.31 | 1.78 | 2.34 |
| Mean | 2.42 | 2.50 | 1.77 | 1.83 | 1.84 | 1.59 | 1.55 | 1.44 | |
| | 2.46 | | 1.80 | | 1.72 | | 1.49 | | |

each subject on each test trial. The .01 level of significance was chosen for evaluation of all effects. The analysis yielded significant main effects for concreteness [$F(1,62) = 13.36$, $MSe = 8.42$] and trials [$F(3,186) = 16.07$, $MSe = 1.38$]. The mean total errors for conditions are given in Table 1.

Inspection of the marginal means in Table 1 indicates significantly fewer errors for CC pairs compared to AA pairs and an overall significant decline in errors over trials. It is evident that the CC pairs were easier than the AA pairs from the first trial on and, although the interaction of concreteness and trials failed to reach significance ($p = .067$), the decline in errors over trials was somewhat greater for the CC pairs compared to the AA pairs. Although no other effects were found to be significant (at the .01 level), the Phenomenal Frequency by Concreteness interaction is given in Table 2 for inspection since it was marginal ($p = .043$) and of some interest.

It should be noted that the means in Table 2 seem to indicate slightly better performance for low phenomenal frequency, but only when the pair members are abstract with the reverse for the concrete pairs. If phenomenal frequency is indeed a powerful variable in VDL and independent of concreteness, pairs low in phenomenal frequency should have proven easier to learn than those high in such frequency regardless of item concreteness. The present results do not support such a contention. The possibility of an interaction between concreteness and phenomenal background frequency did seem to exist, however. Experiment 2 was conducted to explore this possibility further and to test the reliability of the imagery effect when phenomenal frequency is controlled in a mixed-list design.

EXPERIMENT 2

Method

Subjects. Sixty-four subjects drawn from the same population as in Experiment 1 participated in this experiment. These subjects were also naive with respect to verbal learning experiments and received course credit for their participation. Subjects were assigned to conditions in the same manner as in Experiment 1.

Materials. The word materials for this experiment were the same as those employed in Experiment 1 but involved VDL on a mixed list with respect to pair concreteness. Each subject received a 20-pair VDL list having 10 random pairings of C items

and 10 of A items. Four different lists were constructed to investigate the effect of phenomenal background frequency: all pair members high in phenomenal frequency (HH), all low (LL), concrete pairs high and abstract pairs low (HL), and concrete pairs low and abstract pairs high (LH). For each list, a second random pairing of C and A items (yielding 10 CC and 10 AA pairs) was made, thereby generating two replications of word pairings (to control for idiosyncratic pairings as in Experiment 1) with half of the subjects in each of the four frequency conditions receiving one of the replications. Four random study and test orders were again constructed as in Experiment 1 with the same restrictions (except that concreteness of pairs rather than phenomenal frequency was a within-subjects variable). As in Experiment 1, half of the subjects in this experiment received a second version of each list where the correctness of pair members was reversed.

Procedure. The procedure employed in this experiment was essentially the same as that used in Experiment 1. Subjects were again tested individually, receiving four alternating study and test trials with pairs presented via the BCI programmer at the same 3-sec rate. Subjects were instructed to respond during study and test in the same manner as required in Experiment 1. The sequence of study orders was again varied across subjects and the experimenter recorded all responses made during each test trial.

Design. This experiment was a 4 by 2 by 4 mixed design: Pair concreteness (CC or AA) and trials (four) were within-subjects variables. Phenomenal frequency was manipulated in such a manner as to yield four between-subjects list conditions. In two of those conditions, pair member phenomenal frequency was homogeneous, with all pairs (CC or AA) having words high in such frequency (HH) or low (LL). For the other two conditions, phenomenal frequency was intentionally confounded with concreteness to yield a manipulated version of what presumably occurs in nature when phenomenal frequency is not controlled, concrete pairs low and abstract pairs high in such frequency (LH), and a condition exactly opposite (HL).

Results and Discussion

As in Experiment 1, incorrect identifications or errors by each subject on each test trial were subjected to analysis of variance with the .01 level of significance again chosen for evaluation of effects. The mean total errors for all conditions are given in Table 3.

Table 2
Mean Total Errors as a Function of Pair Concreteness and Phenomenal Frequency

| Pair Type | Phenomenal Frequency | | Mean |
|-----------|----------------------|------|------|
| | High | Low | |
| CC | 1.29 | 1.51 | 1.40 |
| AA | 2.50 | 2.17 | 2.34 |
| Mean | 1.90 | 1.84 | |

Table 3
Mean Total Errors as a Function of Phenomenal Frequency of Pairs, Pair Concreteness, and Trials

| Phenomenal Frequency of Pairs (C-A) | Trial 1 | | Trial 2 | | Trial 3 | | Trial 4 | | Mean |
|-------------------------------------|---------|------|---------|------|---------|------|---------|------|------|
| | CC | AA | CC | AA | CC | AA | CC | AA | |
| H-H | 2.69 | 2.81 | 1.44 | 2.56 | 1.19 | 2.88 | .75 | 1.75 | 2.01 |
| L-L | 2.44 | 2.94 | 1.44 | 2.75 | 1.00 | 2.56 | 1.00 | 2.25 | 2.04 |
| L-H | 2.56 | 2.94 | 1.13 | 2.50 | 1.38 | 2.25 | .81 | 1.88 | 1.92 |
| H-L | 2.94 | 3.13 | 1.69 | 2.63 | 1.31 | 1.94 | 1.19 | 2.13 | 2.12 |
| Mean | 2.66 | 2.95 | 1.42 | 2.61 | 1.22 | 2.41 | .94 | 2.00 | |
| | 2.80 | | 2.02 | | 1.81 | | 1.47 | | |

Analysis of the data again yielded significant main effects of concreteness [$F(1,60) = 41.08$, $MSe = 2.72$] and trials [$F(3,180) = 29.20$, $MSe = 1.41$]. Overall, CC pairs were easier (1.56 mean errors) than AA pairs (2.49 mean errors) and performance improved over trials, as can be seen in Table 3. A significant interaction of concreteness and trials was also obtained in this analysis [$F(3,180) = 4.14$, $MSe = 1.42$], which indicated a significantly greater decline in errors over trials for CC pairs than AA pairs. No other effects approached significance (all $F_s < 1$).

Since an interaction of pair concreteness and phenomenal frequency conditions was of interest, the means of this interaction are given in Table 4. As can be seen in this table, performance was consistently better for CC than for AA pairs, regardless of the phenomenal background frequency manipulation. The mean performance for CC and AA pairs remained about the same whether phenomenal background frequency was high (HH) or low (LL), as did the relative difference in performance between CC and AA pairs. The imagery effect was obtained in the LH condition as expected (analogous to phenomenal background frequency situation in nature when not controlled), but also in the HL condition (although somewhat smaller in magnitude), where a reverse imagery effect would be predicted if the typical imagery effect obtained in VDL studies is attributable to phenomenal frequency differences between C and A items.

The results of Experiment 2 failed to provide general support for the conclusions reached by Ghatala and Levin (1976) that phenomenal background frequency is an important variable in VDL and accounts for the

imagery effect typically obtained in this paradigm with concrete and abstract words. There were, however, several potentially important differences between these experiments and that of Ghatala and Levin (1976) that might account for the different results obtained.

Two major differences concern the nature of the word materials employed. The words used in our experiments were high in objective frequency (A-AA Thorndike-Lorge frequency) and were rated for phenomenal frequency on a 9-point scale, whereas Ghatala and Levin (1976) used words rather low in Thorndike-Lorge frequency (1-86) and obtained phenomenal frequency ratings based on a 5-point scale. Our rationale for using words high in objective frequency was fairly straightforward; we wished to control for objective frequency and, since such words have been routinely used in VDL studies where reliable imagery effects have been reported, it seemed reasonable to attempt to demonstrate an effect of phenomenal frequency with such materials prior to attributing the imagery effects obtained to a phenomenal frequency confound. The reasons for our use of a 9-point phenomenal frequency rating scale were twofold. First, words scaled in this manner were readily available from a previous study (Goedel & Thomas, 1977). Second, the original investigation that indicated a potential concreteness and phenomenal frequency confound (Galbraith & Underwood, 1973), upon which Ghatala and Levin (1976) based their hypothesis, used a 9-point scale, and we wished to remain consistent with the original methodology for obtaining such ratings. It should be noted, in reference to this 9-point scale, that the difference between mean high and low phenomenal frequency conditions manipulated in our two experiments paralleled the mean difference reported by Galbraith and Underwood (1973) for the concrete and abstract words rated in their study.

In light of these differences in word materials, it seemed prudent to conduct a third experiment that focused upon the collection of normative data to provide a pool of low objective frequency C and A words rated on a 5-point phenomenal frequency scale for our subject population. The data from this experiment were then used in a fourth experiment where phenomenal frequency was factorially manipulated

Table 4
Mean Total Errors as a Function of Phenomenal Frequency of Pairs and Pair Concreteness

| Frequency of Pair (C-A) | Pair Type | | Mean |
|-------------------------|-----------|------|------|
| | CC | AA | |
| H-H | 1.52 | 2.50 | 2.01 |
| L-L | 1.47 | 2.60 | 2.04 |
| L-H | 1.47 | 2.40 | 1.93 |
| H-L | 1.78 | 2.45 | 2.12 |
| Mean | 1.56 | 2.49 | |

with pair type (homogeneous and heterogeneous with respect to word-pair concreteness in a within-subjects VDL design using an anticipation procedure.

The primary purpose of the last experiment was to evaluate the imagery effect and phenomenal frequency using words of low objective frequency. Levin (Note 1) recently collected data indicating that the effect of phenomenal frequency may be confined to words of low objective frequency in a study utilizing homogeneous abstract word pairs. The present design represented an attempted replication, while also allowing for an evaluation of a possible interaction with concreteness with the addition of homogeneous concrete word pairs.

The inclusion of heterogeneous word pairs (one word concrete and the other abstract) in the design of Experiment 4, although primarily of empirical interest, was not without some theoretical rationale. A plausible alternative explanation for the imagery effect that has not heretofore received serious consideration, but would be consistent with frequency theory, concerns the possibility of greater subjective frequency accrual over trials for A words relative to C words. Evidence bearing on this possibility would include a study by Begg (1974) demonstrating higher frequency judgments for A relative to C words in an absolute situational frequency judgment task (for low presentation frequencies and delayed judgments) and a study by Goedel and Thomas (1977) that indicated better performance in a comparative frequency judgment task when the more frequently occurring item was abstract in heterogeneous pairs.

If subjective evaluation of frequency is indeed higher for A words and subjective intrapair frequency differences used as the basis for learning in the VDL paradigm, the imagery effect could be handled by frequency theory by appealing to the Weber's law postulate. This postulate would predict greater difficulty for homogeneous abstract pairs relative to concrete pairs due to more difficult relative subjective frequency discriminations. In addition, with respect to the heterogeneous pair type conditions included in Experiment 4, this postulate would predict fewest errors when the abstract member of a heterogeneous pair was correct and most errors when the abstract member was incorrect.

EXPERIMENT 3

Method

Subjects. One hundred and fifty-two subjects from the same student population as in Experiments 1 and 2 participated in this experiment. These subjects had not participated in verbal learning experiments previously and received course credit for their participation in the present experiment.

Materials. A total of 200 C and 200 A nouns were selected from the Paivio, Yuille, and Madigan (1968) norms. The two sets of words were matched as closely as possible on Thorndike-Lorge frequency (range: 0-AA) and meaningfulness (range:

3.00-8.76). The average concreteness ratings for the C and A words were 6.58 (range: 5.51-7.00) and 2.18 (range: 1.18-3.00), respectively. These two sets of words were then randomly sorted into four groups of 100 (50 C and 50 A per group) for subjects to rate. Each subject rated words contained in only one of the four groups.

Procedure. Each of the four groups of 100 words was randomly ordered in a two-page booklet with two columns of 25 words per page, with the restriction that 25 C and 25 A words appear per page. Each booklet also included a cover sheet giving printed instructions concerning the rating task and described the five points of the phenomenal frequency scale following the Ghatala and Levin (1976, Experiment 1) procedure. Also included were the four anchor words with appropriate 1 or 5 scale values (water-5, figment-1, belfry-1, thought-5) used by these authors as well as Galbraith and Underwood (1973, Experiment 3). The subjects performed the rating task in small groups (20-30 subjects) with the four different sets of word booklets distributed as randomly and evenly as possible. The experimenter read the cover sheet aloud to the subjects and responded to any questions prior to the start of the experiment, to insure understanding of the instructions. The scaling task was unpaced.

Results and Discussion

The average scale values and standard deviations for each of the 400 words were calculated with $n = 38$ for each word scaled. Separate distributions for the C and A words were then examined and are plotted in Figure 1 in half-scale steps. Inspection of these distributions indicates considerable overlap in phenomenal frequency in contrast to that obtained by Ghatala and Levin (1976) for 53 C and 53 A words ($n = 40$) using a 5-point scale, as well as that reported by Galbraith and Underwood (1973) for 73 C and 73 A words ($n = 100$) using a 9-point scale.

The overall means obtained for the 200 A and 200 C words scaled in the present study were 3.11 ($SD = .84$) and 3.03 ($SD = .87$), respectively. However, a 2 by 4 mixed analysis of variance (concreteness by groups) performed on these data indicated that this difference was significant [$F(1,148) = 7.88, p < .01$]. The interaction of Concreteness by Groups was also

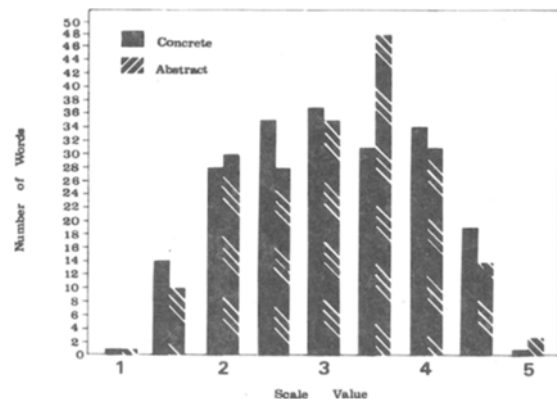


Figure 1. Number of concrete and abstract words (200 each) assigned to each half-step value on the 5-point background frequency scale.

Table 5
Mean Phenomenal Background Frequency Ratings (PBF)
for the 50 Concrete and 50 Abstract Words
Assigned to Each of Four Groups

| Group | Concrete Words | | | Abstract Words | | |
|-------|----------------|------|------|----------------|------|------|
| | T-L | m | PBF | T-L | m | PBF |
| 1 | 28.63 | 6.59 | 3.00 | 26.68 | 5.06 | 3.08 |
| 2 | 33.42 | 6.47 | 3.01 | 28.74 | 5.16 | 2.92 |
| 3 | 34.02 | 6.45 | 3.01 | 34.04 | 5.08 | 3.27 |
| 4 | 32.22 | 6.45 | 3.08 | 32.08 | 5.19 | 3.17 |
| Mean | 32.07 | 6.49 | 3.03 | 30.14 | 5.12 | 3.11 |

Note—Also given are the mean objective frequency (T-L) and meaningfulness (m) values for the words assigned to each group.

significant [$F(3,148) = 6.14, p < .01$], and the means are given in Table 5. Overall, the interaction indicates differences in the magnitude of the difference in mean phenomenal frequency given to the C and A words assigned to the four groups (replications) with one slight reversal (C words rated higher than A for Group 2). This effect may be due to the fact that the C and A words in the present study were randomly assigned to groups without regard to Thorndike-Lorge frequency and meaningfulness and would suggest that one or both of these variables may have some influence on phenomenal frequency ratings of C and A words. This possibility was evaluated via post hoc calculation of mean meaningfulness and Thorndike-Lorge frequency for the C and A words assigned to each group. These data are also given in Table 5. Inspection of average Thorndike-Lorge frequency for both types of words in each group would appear to explain the interaction. The group with the greatest Thorndike-Lorge frequency difference between C and A words (Group 2) yielded the unexpected reversal in phenomenal frequency ratings for these items (C rated higher), while the two groups with small average Thorndike-Lorge differences (Groups 3 and 4) yielded the largest phenomenal frequency differences in the expected direction (A rated higher).

Nevertheless, the significant main effect of concreteness in the present experiment replicates the findings of previous investigations of phenomenal background frequency despite the small descriptive mean difference obtained. This provides additional evidence that abstract words tend to be perceived by college student subjects as being higher, on the average, in phenomenal frequency than concrete words.

EXPERIMENT 4

Method

Subjects and Design. A total of 60 student volunteers from the same population as the previous experiments were employed. All of these subjects were naive with respect to verbal learning experiments and all received course credit for their participation. A mixed 2 by 2 by 4 by 4 design was employed with a single between-subjects factor. This factor was replications and simply consisted of two, 32-pair VDL lists reflecting different word

pairings. The major factors of interest were manipulated within subjects and included phenomenal background frequency of pair members (high or low), concreteness of pair members (both concrete, CC; both abstract, AA; correct member concrete and incorrect member abstract, CA; and correct member abstract and incorrect member concrete, AC) and anticipation trials (four following the initial practice or guessing trial).

Materials. A total of 32 C and 32 A words were selected from the word pool provided by Experiment 3. All words were low in objective Thorndike-Lorge frequency with half of the C and A words rated high (H) and half low (L) in phenomenal frequency. The factorial combination of concreteness and phenomenal frequency yielded four groups of 16 words (CH, CL, AH, and AL). The mean concreteness, phenomenal frequency, and Thorndike-Lorge frequency values (and ranges) for these groups were as follows: CH, 6.44 (5.66-6.96), 3.29 (2.76-3.92), and 4.50 (1-20); CL, 6.52 (5.58-6.96), 1.70 (1.26-1.92), and 4.56 (1-26); AH, 2.33 (1.73-2.98), 3.31 (2.78-3.94), and 4.50 (1-21); and AL, 2.36 (1.73-2.95), 1.74 (1.18-2.00), and 4.56 (0-29).

In order to construct the 32 pairs of words used in the VDL task, the following procedure was employed. First, eight words in each of the four groups were randomly designated as correct with the remainder incorrect. Word pairs were then randomly determined by pairing correct and incorrect items with the restriction that eight concreteness by phenomenal frequency conditions result (CC-H, CC-L, AA-H, AA-L, CA-H, CA-L, AC-H, AC-L), with four pairs in each condition. This procedure was repeated a second time to produce a second list of different word pairings (a replications factor), with half of the subjects in the study receiving one of these lists. In addition, correct and incorrect item designations were reversed following the construction of each list for half of the subjects receiving each list.

Five presentation orders for each of the two lists were constructed such that (1) within each order, the eight concreteness by phenomenal frequency pair types were block randomized; (2) within an order, the correct items were located on the right for half of the pairs of each pair type and on the left for the other half; and (3) for a random half of the pairs, the spatial location of the correct item changed between orders with the restriction that, for all pairs, the spatial location of the correct item change at least twice across the five orders. The initial order of presentation was randomly changed so that one-fifth of the subjects receiving each list began with one of the five different orders.

Procedure. The procedure was similar to that used in Experiments 1 and 2. Each subject was tested individually and sat across from the experimenter in front of a BCI programmer (Model SR-400). Word pairs were presented one at a time at a constant rate of 3 sec, each using an anticipation procedure. Each pair was presented initially for 3 sec, with the subject instructed to pronounce each word aloud and then to respond aloud with what he/she felt was the correct member of the pair. The same pair was then presented again for 3 sec, with the correct member of the pair underlined. Each subject received five anticipation trials on one of the two lists of 32 word pairs with the first trial being a study or guessing trial. The intertrial interval was also 3 sec and was represented on the list by a row of asterisks. Conventional VDL instructions for the anticipation procedure were read to each subject prior to the start of the experiment. The experimenter recorded each response made by each subject on each of the five trials.

Results and Discussion

The data subjected to analysis were the number of incorrect identifications (errors) made by each subject on each trial. An initial analysis of variance was performed on the data from the first or guessing trial

of each subject to evaluate possible response bias (especially for pairs heterogeneous with respect to concreteness), with the .01 level of significance chosen for evaluation of all effects. This was a 2 by 2 by 4 mixed analysis with the two lists (replications) as the between-subjects factor and pair type (CC, AA, CA, and AC) and phenomenal frequency of pairs (H or L) as within-subjects factors. This analysis failed to yield any significant effects and thus provided no evidence for the operation of a possible response bias during VDL. Although no effects were expected, the initial analysis seemed warranted prior to the evaluation of any effects involving the heterogeneous word pairs.

The second and main analysis was also a mixed design with one between-subjects factor (lists) and three within-subjects factors, pair type (CC, AA, CA, and AC), phenomenal frequency (H or L), and trials (four trials following the initial guessing trial). This analysis resulted in significant main effects for pair type [$F(3,174) = 11.07$, $MSe = .96$], and trials [$F(3,174) = 157.32$, $MSe = .65$], and a significant interaction of these two factors [$F(9,522) = 2.52$, $MSe = .56$] (all $ps < .01$). No other effects approached significance (all $ps > .05$).

As can be seen in Table 6, overall performance improved in a fairly linear fashion over trials on each of the four pair types. Collapsed over trials, the best performance was obtained on the CC pairs and the worst on the AA pairs, with performance on the heterogeneous pairs (CA and AC) falling between these two conditions. Tukey comparisons on these means showed the CC and AA conditions to be significantly different from each other and each to be significantly different from the heterogeneous conditions ($p < .01$). These results provide strong support for the reliability of the imagery effect in VDL. Little support is provided, however, for a frequency theory interpretation suggested earlier, that perhaps subjective frequency accrual for A items exceeds that for C items during VDL to produce the effect. The CA and AC conditions do not differ from each other, nor do they produce the respective maximum and minimum number of errors as would be predicted by such a hypothesis. In general, the interaction reflects differences in rate of learning for the four pair types

Table 6
Mean Total Errors as a Function of Pair Type
for Each of Four VDL Trials

| Pair Type | Trial | | | | Mean |
|-----------|-------|------|-----|-----|------|
| | 1 | 2 | 3 | 4 | |
| CC | 1.47 | .84 | .58 | .33 | .81 |
| AA | 1.56 | 1.42 | .98 | .64 | 1.15 |
| CA | 1.75 | 1.26 | .76 | .51 | 1.07 |
| AC | 1.52 | 1.28 | .83 | .63 | 1.06 |
| Mean | 1.58 | 1.20 | .79 | .53 | |

Table 7
Mean Total Errors as a Function of Pair Type and
Phenomenal Background Frequency (PBF)

| PBF | Pair Type | | | | Mean |
|------|-----------|------|------|------|------|
| | CC | AA | CA | AC | |
| High | .78 | 1.22 | 1.03 | .99 | 1.00 |
| Low | .84 | 1.08 | 1.11 | 1.13 | 1.04 |
| Mean | .81 | 1.15 | 1.07 | 1.06 | |

with somewhat faster learning on the CC and CA pair types relative to the AA and AC types.

Since an effect of phenomenal frequency and a possible interaction of this variable with pair type was of considerable interest, these means are given in Table 7 for inspection. The lack of a significant main effect of phenomenal frequency or interaction with pair types tested fails to provide any evidence for the involvement of this factor in the imagery effect typically found in VDL. In addition, the minor mean differences that can be noted in Table 7 are in the opposite direction (mean errors higher when both pair members are low in phenomenal frequency) for three of the pair types tested (CC, AC, and CA) from that which would be predicted based upon Ghatala and Levin's (1977) confounding hypothesis. The only difference (and largest) that directionally fits the hypothesis occurred for AA pairs. Since Levin (Note 1) has recently found evidence for a phenomenal frequency effect in a VDL list composed solely of AA pairs, a *t* test for paired (dependent) data was performed using only the raw error data for the AA pairs in the present study collapsed over trials. This analysis also failed to provide significant support for a phenomenal frequency effect [$t(59) = 1.20$, $p > .05$]. Thus, although our mean error difference for AA pairs appears descriptively to be in agreement with Levin's data, it is not statistically reliable, suggesting that such an effect may be limited to AA pairs in an unmixed list.

CONCLUSIONS

In general, the data of the present study demonstrate rather strong support for the reliability and robustness of the imagery effect in VDL. Performance on word pairs with both members concrete was far superior to performance on pairs with both members abstract. This effect was obtained for mixed and unmixed lists with words high in objective background frequency using a study-test presentation method (Experiments 1 and 2), as well as for a mixed list with words low in objective background frequency using an anticipation method of presentation (Experiment 4), regardless of phenomenal background frequency manipulation.

In none of these experiments (1, 2, or 4) was there any evidence that phenomenal frequency plays a major role in VDL. In addition, although the replicable finding

that abstract words are, in general, rated higher than concrete words in phenomenal background frequency by college student subjects (Experiment 4) readily suggests an explanation of the imagery effect in terms of this apparent confound as noted by Ghatala and Levin (1976), direct factorial manipulation of phenomenal frequency and concreteness in the present study failed to support this possibility.

Although we are somewhat unwilling to reject the rather attractive "confound hypothesis" initially proposed by Ghatala and Levin (1976) based upon our failure to obtain a significant effect for phenomenal background frequency (since this would constitute an acceptance of the null), we feel compelled by our data to seriously question the reliability of such an effect or, at the very least, to suggest that the effect may be confined to rather limited experimental conditions. Since Levin (Note 1) has recently obtained such an effect for low objective frequency words with homogeneous abstract word pairs in an unmixed-list design, we would tend to conclude the latter. As such, the usefulness of phenomenal frequency as a viable frequency theory explanation for the imagery effect is questionable.

A plausible alternative explanation that would still appeal to the Weber's law postulate of frequency theory suggests possible differential subjective situational frequency accrual for abstract and concrete items during VDL, rather than preexperimental phenomenal background frequency differences. However, this hypothesis was evaluated in Experiment 4 and empirical support was found to be lacking, thereby rendering such an alternative equally questionable.

Thus, it would appear that the dilemma posed for frequency theory by the imagery effect in VDL remains to be resolved. One obvious resolution would be to accept the fact that attributes other than frequency may predominate as the basis for learning in the verbal discrimination task, especially when the dimension of concreteness is salient. In light of frequency discrimination studies (Goedel & Thomas, 1977; Wallace, Murphy, & Sawyer, 1973), however, that have shown relative frequency discrimination to be easier for homogeneous concrete pairs relative to homogeneous abstract pairs, a reluctance to accept such a resolution remains.

REFERENCE NOTE

1. Levin, J. R., Ghatala, E. S., & Truman, D. L. *Phenomenal frequency effects in verbal discrimination learning: resolving a discrepancy*. Manuscript in preparation, 1978.

REFERENCES

- BEGG, I. Estimation of word frequency in continuous and discrete tasks. *Journal of Experimental Psychology*, 1974, 102, 1046-1052.
- ECKERT, E., & KANAK, N. J. Verbal discrimination learning: A review of the acquisition, transfer and retention literature through 1972. *Psychological Bulletin*, 1974, 81, 582-607.
- EKSTRAND, B. R., WALLACE, W. P., & UNDERWOOD, B. J. A frequency theory of verbal discrimination learning. *Psychological Review*, 1966, 73, 566-578.
- GALBRAITH, R. C., & UNDERWOOD, B. J. Perceived frequency of concrete and abstract words. *Memory & Cognition*, 1973, 1, 56-60.
- GHATALA, E. S., & LEVIN, J. R. Phenomenal background frequency and the concreteness/imagery effect in verbal discrimination learning. *Memory & Cognition*, 1976, 4, 302-306.
- GHATALA, E. S., LEVIN, J. R., & MAKOID, L. A. A clarification of frequency effects in children's discrimination learning. *Memory & Cognition*, 1975, 3, 1-6.
- GOEDEL, G. D., & THOMAS, C. A. The imagery effect and frequency discrimination. *Memory & Cognition*, 1977, 5, 53-58.
- KUČERA, H., & FRANCIS, W. N. *Computational analysis of present-day American English*. Providence: Brown University Press, 1967.
- PAIVIO, A., & ROWE, E. J. Noun imagery, frequency and meaningfulness in verbal discrimination. *Journal of Experimental Psychology*, 1970, 85, 264-269.
- PAIVIO, A., & ROWE, E. J. Intrapair imagery effects in verbal discrimination and incidental associative learning. *Canadian Journal of Psychology*, 1971, 25, 302-312.
- PAIVIO, A., YUILLE, J. C., & MADIGAN, S. Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, 1968, 76(1, Part 2).
- ROWE, E. J., & PAIVIO, A. Word frequency and imagery effects in verbal discrimination learning. *Journal of Experimental Psychology*, 1971, 88, 319-326.
- ROWE, E. J., & PAIVIO, A. Effects of noun imagery, pronunciation, method of presentation, and intrapair order of items in verbal discrimination learning. *Journal of Experimental Psychology*, 1972, 93, 427-429.
- SPREEN, O., & SCHULZ, R. W. Parameters of abstraction, meaningfulness, and pronunciability for 329 nouns. *Journal of Verbal Learning and Verbal Behavior*, 1966, 5, 459-468.
- THORNDIKE, E. L., & LORGE, I. *The teacher's word book of 30,000 words*. New York: Bureau of Publications, Teachers College, Columbia University, 1944.
- ULLRICH, J. R., & BALOGH, B. A. Imagery and meaningfulness of right and wrong items in verbal discrimination learning. *Psychonomic Science*, 1972, 29, 68-70.
- WALLACE, W. P. Verbal discrimination. In C. P. Duncan, L. Sechrest, & A. W. Melton (Eds.), *Human memory: Festschrift for Benton J. Underwood*. New York: Appleton-Century-Crofts, 1972.
- WALLACE, W. P., MURPHY, M. D., & SAWYER, T. J. Imagery and frequency in verbal discrimination learning. *Journal of Experimental Psychology*, 1973, 101, 201-219. (Monograph)

(Received for publication December 5, 1977;
accepted December 12, 1977.)